

PRICING

PRICING

AND

REVENUE OPTIMIZATION

PRICING
AND
REVENUE OPTIMIZATION

ROBERT L. PHILLIPS

STANFORD BUSINESS BOOKS
An imprint of Stanford University Press
Stanford, California 2005

Stanford University Press
Stanford, California

© 2005 by the Board of Trustees of the Leland Stanford Junior University. All rights reserved.

No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or in any information storage or retrieval system without the prior written permission of Stanford University Press.

Printed in the United States of America on acid-free, archival-quality paper

Library of Congress Cataloging-in-Publication Data

Phillips, Robert L. (Robert Lewis), 1955—

Pricing and revenue optimization / Robert L. Phillips.

p. cm.

Includes bibliographical references and index.

ISBN 0-8047-4698-2 (cloth : alk. paper)

1. Pricing. 2. Revenue management. I. Title.

HF5416.5.P457 2005

658.15'54—dc22

2005009126

Typeset by G&S Book Services in 10/13.5 Minion

Original Printing 2005

Last figure below indicates year of this printing:

14 13 12 11 10 09 08 07 06 05

CONTENTS

	<i>Preface</i>	xi
<i>Chapter 1</i>	Background and Introduction	1
	1.1 Historical Background and Context	2
	1.2 The Financial Impact of Pricing and Revenue Optimization	13
	1.3 Organization of the Book	14
<i>Chapter 2</i>	Introduction to Pricing and Revenue Optimization	18
	2.1 The Challenges of Pricing	18
	2.2 Traditional Approaches to Pricing	22
	2.3 The Scope of Pricing and Revenue Optimization	26
	2.4 The Pricing and Revenue Optimization Process	29
	2.5 Summary	35
	2.6 Exercises	36
<i>Chapter 3</i>	Basic Price Optimization	38
	3.1 The Price-Response Function	38
	3.2 Price Response with Competition	55
	3.3 Incremental Costs	59
	3.4 The Basic Price Optimization Problem	61
	3.5 Summary and Extensions	69
	3.6 Exercises	70
<i>Chapter 4</i>	Price Differentiation	74
	4.1 The Economics of Price Differentiation	75
	4.2 Limits to Price Differentiation	77
	4.3 Tactics for Price Differentiation	78
	4.4 Volume Discounts	86
	4.5 Calculating Differentiated Prices	89

	4.6 Price Differentiation and Consumer Welfare	93
	4.7 Summary	96
	4.8 Exercises	96
<i>Chapter 5</i>	Pricing with Constrained Supply	99
	5.1 The Nature of Supply Constraints	100
	5.2 Optimal Pricing with a Supply Constraint	101
	5.3 Opportunity Cost	103
	5.4 Market Segmentation and Supply Constraints	104
	5.5 Variable Pricing	106
	5.6 Variable Pricing in Action	111
	5.7 Summary	116
	5.8 Exercises	117
<i>Chapter 6</i>	Revenue Management	120
	6.1 History	121
	6.2 Levels of Revenue Management	123
	6.3 Revenue Management Strategy	123
	6.4 The System Context	125
	6.5 Booking Control	126
	6.6 Tactical Revenue Management	131
	6.7 Net Contribution in Revenue Management	136
	6.8 Measuring Revenue Management Effectiveness	140
	6.9 Revenue Management in Action	141
	6.10 Summary	144
	6.11 Exercise	146
<i>Chapter 7</i>	Capacity Allocation	149
	7.1 Introduction	149
	7.2 The Two-Class Problem	149
	7.3 Capacity Allocation with Multiple Fare Classes	158
	7.4 Capacity Allocation with Dependent Demands	165
	7.5 Capacity Allocation in Action	169
	7.6 Measuring Capacity Allocation Effectiveness	170
	7.7 Summary	172
	7.8 Exercises	173
<i>Chapter 8</i>	Network Management	176
	8.1 Background and Introduction	176
	8.2 When is Network Management Applicable?	177
	8.3 A Linear Programming Approach	183
	8.4 Virtual Nesting	189
	8.5 Network Bid Pricing	195

8.6	Dynamic Virtual Nesting	202
8.7	Network Management in Action	202
8.8	Summary	204
8.9	Exercises	205
<i>Chapter 9</i>	Overbooking	207
9.1	Introduction	207
9.2	A Model of Customer Bookings	212
9.3	Solution Approaches	213
9.4	Extensions	228
9.5	Measuring and Managing Overbooking	233
9.6	Alternatives to Overbooking	235
9.7	*Appendix: Derivation of the Show Probability for Overbooking	237
9.8	Exercises	238
<i>Chapter 10</i>	Markdown Management	240
10.1	Background	241
10.2	Markdown Optimization	249
10.3	Estimating Markdown Sensitivity	256
10.4	Markdown Management in Action	258
10.5	Summary	261
10.6	Exercises	261
<i>Chapter 11</i>	Customized Pricing	264
11.1	Background and Business Setting	264
11.2	Calculating Optimal Customized Prices	269
11.3	Bid Response	277
11.4	Extensions and Variations	290
11.5	Customized Pricing in Action	297
11.6	Exercise	299
<i>Chapter 12</i>	Pricing and Revenue Optimization and Customer Acceptance	301
12.1	Price Presentation and Framing	304
12.2	Fairness	309
12.3	Implications for Pricing and Revenue Optimization	315
12.4	Summary	321
12.5	Exercise	321
<i>Appendix A</i>	Optimization	323
A.1	Continuous Optimization	323
A.2	Linear Programming	324
A.3	Duality and Complementary Slackness	325
A.4	Discrete Optimization	326

<i>Appendix B</i>	Probability	327
	B.1 Probability Distributions	327
	B.2 Continuous Distributions	330
	B.3 Discrete Distributions	331
	 <i>Bibliography</i>	 337
	<i>Index</i>	345

1

BACKGROUND AND INTRODUCTION

What is a cynic? . . . A man who knows the price of everything and the value of nothing

Oscar Wilde (1892)

This is a book about pricing—specifically, how companies should set and adjust their prices in order to maximize profitability. It takes the view that pricing decisions are commonplace, that they can be complex, and that they are usually critical determinants of profitability. Despite this, pricing decisions are often badly managed (or even unmanaged). While most companies have pretty good prices in place most of the time, very few have the processes and capabilities needed to ensure that they have the right prices in place for all their products, to all their customers, through all their channels, all the time. This is the goal of pricing and revenue optimization.

Pricing and revenue optimization is a tactical function. It recognizes that prices need to change rapidly and often and provides guidance on how they should change. This makes it distinct from strategic pricing, where the goal is usually to establish a general position within a marketplace. While strategic pricing worries about how a product should in general be priced relative to the market, pricing and revenue optimization is concerned with determining the prices that will be in place tomorrow and next week. Strategic pricing sets the constraints within which pricing and revenue optimization operates.

One of the distinguishing characteristics of pricing and revenue optimization is its use of analytical techniques derived from management science. The use of these techniques to set prices in a complex, dynamic environment is relatively new. One of the first applications of the approach was the development of revenue management systems by the passenger airlines in the 1980s. Since then, the rapid development of e-commerce and the availability of customer data through customer relationship management (CRM) systems has led to the adoption of similar techniques in many other industries, including automotive, retail, telecommunications, financial services, and manufacturing. A number of software vendors provide “price optimization” or “demand management” or “revenue management” solutions focused on one or more industries. In this context, pricing and revenue optimization is increasingly becoming a core competency within many different companies. The purpose of this book is to provide an introduction to this relatively new and rapidly changing field.

In this chapter, we begin by giving some historical context for pricing and revenue opti-

mization. In particular, we give some perspective on why pricing has gone from being a largely ignored and obscure “black art” within many companies to become the subject of intense scrutiny and analysis. We argue that the “pricing problem” has become increasingly difficult and is likely to become even more difficult in the future. We argue further that improving pricing is often one of the highest-return investments available to a company. Hopefully this will whet the reader’s appetite for the more quantitative material to come in the following chapters.

1.1 HISTORICAL BACKGROUND AND CONTEXT

For most of history, philosophers took it for granted that goods had an intrinsic value in the same sense that they had an intrinsic color or weight. A fair price reflected that intrinsic value. Charging a price too much in excess of the intrinsic value was condemned as a sign of “avarice” and often prohibited by law. Prices were set by custom, by law, or by imperial fiat. Sermons were preached inveighing against the sin of charging unfair prices in order to receive excessive profits.¹ The problem of pricing did not really exist until modern market economies began to emerge in the West in the 17th and 18th centuries. With the emergence of these economies, prices were allowed to move more freely—untied to the traditional concept of value. Speculative bubbles such as “tulipomania” in the Dutch republic in the 1630s—in which the price of some varieties of tulips rose more than a hundredfold in 18 months before collapsing in 1637—and the “South Sea bubble” in England in 1720—in which the prices of shares in the South Sea Company soared before the company collapsed amid general scandal—fed a sense of anxiety and the belief that prices could somehow lose touch with reality.² Furthermore, for the first time, large numbers of people could amass fortunes—and lose them—by buying and selling goods on the market. The question naturally arose—what were prices, exactly? Where did they come from? What determined the right price? When was a price fair? When should the government intervene in pricing? The modern field of economics arose, at least in part, in response to these questions.

Possibly the greatest insight of classical economics was that the price of a good at any time in an ideal capitalist economy is not based on any intrinsic “value” but rather on the interplay of supply and demand. This was a major intellectual breakthrough—on par in its way with the Newtonian view of the clockwork universe and Darwin’s theory of evolution. In essence, the price of a good or service was determined by the interaction of people willing to sell the good with the willingness of others to buy the good. That’s all there is to it—neither intrinsic “value” nor cost nor labor content enters directly into the equation. Of course, these and other factors enter *indirectly* into pricing—sellers would not last long selling goods below cost, and the prices buyers accept are based on the “value” they placed on the item—but these were not primary. There are many reasons why sellers sell below cost when they are in possession of a cartload of vegetables that are on the verge of going rotten—the classic “sell it or smell it” situation—or to attract a desirable new customer. Just so, the “value” that buyers placed on different goods changed with their changing situation and the dictates of fashion. According to modern economics there is no normative “right price” for a good or service against which the price can be compared—rather, there are only the

actual prices out in the marketplace, floating freely without an anchor, based only on the willingness of sellers to sell and buyers to buy.

While classical economics solved the problem of the origin of price, it raised as many questions as it answered. In particular, if prices were not tied to fundamental values—if they had no anchor—why did they show any stability at all? Under normal circumstances, prices for most goods are pretty stable most of the time. If prices are based only on the whims of buyers and sellers, why is the price of bread not subject to wild swings like the Dutch tulip market in 1689? Why doesn't milk cost five times as much in Chicago as it does in New York? How can manufacturers and merchants plan at all and make reasonable profits in order to stay in business? How can an economy based on free-floating prices work at all? And, assuming that such an economy could work, how could it possibly work better than a centralized economy where planners carefully sought to allocate resources across the entire economy?

One of the great achievements of 20th century economics was to show mathematically how a largely unregulated economy could work: that an economy consisting of individuals who supply their labor in return for wages and use their earnings to buy goods to maximize their "utility" combined with firms who seek to maximize profitability can be remarkably stable and efficient.³ Under certain assumptions, this type of capitalist economy can be shown to be at least as efficient as any centrally planned economy. Furthermore, prices in such an economy would generally be stable and reasonably predictable. The price for a product would equal the long-run marginal production cost of that product plus the return on invested capital necessary to produce the product. If someone were selling the product for less, he or she would go out of business because his or her costs would not be covered. If someone tried selling for more, other sellers would undercut his price, consumers would flee to the lower-priced sellers, and the high-price seller would be forced to lower his price or go bankrupt for lack of business. As this happens simultaneously, economy-wide, prices equilibrate and change only due to exogenous shocks, changes in resource availability, taxation or monetary policy, or changes in consumer tastes.

This view of the world is based primarily on the assumption that most markets are perfectly competitive, where the idea of perfect competition can be summarized as follows.

A market structure is perfectly competitive if the following conditions hold: There are many firms, each with an insubstantial share of the market. These firms produce a homogenous product using identical production processes and possess perfect information. It is also the case that there is free entry to the industry; that is, new firms can and will enter the industry if they observe that greater-than-normal profits are being earned. The effect of this free entry is to push the demand curve facing each firm downwards until each firm earns only normal profits, at which point there is no further incentive for new entrants to come into the industry. Moreover, since each firm produces a homogenous product, it cannot raise its price without losing all of its market to its competitors Thus firms are price takers and can sell as much as they are capable of producing at the prevailing market price.⁴

There are no pricing decisions in perfectly competitive markets—prices are determined by the iron law of the market. If one merchant were offering a good for a lower price than another, neoclassical economics assumes that either customers would entirely abandon the

higher-price merchant and swamp the lower-price merchant or an arbitrageur would arise who would buy all the goods from the lower-price merchant and sell them at the higher price. In either case, a single market price would prevail. Furthermore, if prices were so high that merchants enjoyed higher profits than the rest of the economy, more sellers would enter, lowering the average price until the return on capital dropped to the market level. In this situation, there are no pricing decisions at all: Prices are set “by the market”—as stock prices are set by the New York Stock Exchange or NASDAQ. The price of Microsoft stock is not set by any “pricer” but by the interplay of supply and demand for the stock. Many financial instruments, such as stocks and bonds, satisfy the economic definition of a commodity. Certain other highly fungible goods—grain, crude oil, and some bulk chemicals—also come very close to being commodities. In these markets, there is simply no need for pricing and revenue optimization—the market truly sets the price.

As any shopper can tell you, much of the real world is messier—prices vary all over the place, sometimes in ways that seem irrational. Buyers often behave erratically, sellers do not always seek to maximize short-run profit, neither buyers nor sellers are possessed of perfect information, and opportunities for arbitrage are not immediately seized. Table 1.1 shows prices for a half gallon of whole milk at different markets in a 16-block area of the upper west side of Manhattan on a single day in May 2002. Prices range from a low of \$1.39 to a high of \$2.00—a variation of \$0.61, or 44%. Furthermore, the price varied by more than \$0.40 even for two stores on the same block. How could this be? Why would anybody buy milk at a high price when they could walk a block and save 40 cents? Why don’t arbitrageurs buy all the milk at the lower price and sell it at the higher?

TABLE 1.1
*Retail prices for a half gallon of whole milk on
the Upper West Side of Manhattan, May 2002*

Location	Price
74th and Broadway	\$1.39
79th and Amsterdam	1.59
77th and Broadway	1.59
74th and Columbus	1.69
73rd and Columbus	1.79
74th and Amsterdam	1.79
75th and Broadway	1.89
71st and Columbus	1.99
78th and Amsterdam	2.00
AVERAGE	\$1.75
STANDARD DEVIATION	0.20

The price variation shown in Table 1.1 will hardly come as a shock to most people—after all, both businesses and consumers know that it pays to shop around—suppliers of the same (or similar) products will often charge different prices. Furthermore, there are other ways to pay a lower price for exactly the same product: Wait until it goes on sale, travel to a retail outlet, clip a coupon, buy in bulk, buy it online, try to negotiate a lower price. In fact, it is hardly a secret not only that prices vary between sellers but that a single seller will often sell the same product to different customers for different prices!

The tools that pricers use day to day are far more likely to be drawn from the fields of statistics or operations research than from economics. Marketing science, which deals with the quantitative analysis of marketing initiatives, including pricing, is usually considered part of the broader field of operations research and management science.⁵ Application of these techniques to problems of marketing began to emerge in a significant fashion in the 1960s.⁶ Since then, marketing scientists have developed, applied, validated, and refined important mathematical models to a broad range of issues, such as forecasting sales, product planning, predicting market response, product positioning, pricing, promotions, sales force compensation, and marketing strategy.⁷ True to its name, marketing science has brought some science to what was previously viewed as a “black art.”

Despite these achievements, there remains a gap between marketing science models and their use in practice. The reasons for this gap are numerous. Many marketing models have been built on unrealistically stylized views of consumer behavior. Other models have been built to “determine if what we see in practice can happen in theory.” Other models seem limited by unrealistically simplistic assumptions. In any case, a leading text on marketing science admitted:

With an area of such importance and with so much at stake, it might be assumed that a great deal of continuing research and planning would by now underlie the formulation of pricing strategy and the setting of prices. One might also expect that a well-developed body of theory would have resulted in principles to guide pricing decisions. But this does not appear to be the case.⁸

One of the possible reasons for the gap between marketing science theory and its application to real pricing decisions is that pricing decisions are becoming increasingly tactical and operational in nature. Companies increasingly need to make pricing decisions more and more rapidly in order to respond to competitive actions, market changes, or their own inventory situation. They no longer have the luxury to perform market analyses or extended spreadsheet studies every time a pricing change needs to be considered. The premium is on speed. While there has been a general acceleration of business in all fields, the impact on pricing and revenue optimization has been particularly notable. This acceleration—and the corresponding interest in developing tools to enable better pricing and revenue optimization (PRO) decisions—has been driven by four trends.

- The success of revenue management in the airline industry provided an example of how pricing and revenue optimization could increase profitability in a real-time pricing environment.
- The widespread adoption of enterprise resource planning (ERP) and customer relationship management (CRM) software systems provided a new wealth of corporate information that can be utilized to improve pricing and revenue optimization decisions.
- The rise of e-commerce necessitated the ability to manage and update prices in a fast-moving, highly transparent, online environment for many companies that had not previously faced such a challenge.

- The success of supply chain management proved that analytic software systems could drive real business improvements.

Because of their importance to the development of pricing and revenue optimization, we will spend a little time to discuss each of these trends.

1.1.1 The Success of Revenue Management

In 1985, American Airlines was threatened on its core routes by the low-fare carrier PeopleExpress. In response, American developed a revenue management program based on differentiating prices between leisure and business travelers. A key element of this program was a “yield management” system that used optimization algorithms to determine the right number of seats to protect for later-booking full-fare passengers on each flight while still accepting early-booking low-fare passengers. This approach was a resounding success for American, resulting ultimately in the demise of PeopleExpress.

We delve more deeply into the American Airlines/PeopleExpress story in Chapter 6. For now, the importance of the story is in the publicity it garnered. American Airlines featured its revenue management capabilities in its annual report. The team that developed the system won the 1991 Edelman Prize for best application of management science. American Airlines’ revenue management has been widely touted as an important strategic application of management science (C. K. Anderson, Bell, and Kaiser 2003), and the tale of American using its superior capabilities to defeat PeopleExpress was the centerpiece of a popular business book (Cross 1997).

Not surprisingly this publicity resulted in widespread interest. Companies began to investigate the prospects of improving the profitability of their pricing decisions. Ford Motor Company was inspired by the success of revenue management at the airlines to institute its own very successful program (Leibs 2000). Vendors arose selling revenue management software systems, and consultants appeared offering to help companies set up their own programs. Revenue management spread well beyond the passenger airlines.

Under its strictest definition, revenue management has a fairly narrow field of application. In particular, the techniques of revenue management are applicable when the following conditions are met.

- Capacity is limited and immediately perishable. Most obviously, an empty seat on a departing aircraft or an empty hotel room cannot be stored to satisfy future demand.
- Customers book capacity ahead of time. Advance bookings are common in industries with constrained and perishable capacity, since customers need a way to ensure ahead of time that capacity will be available when they need to consume it. This gives airlines the opportunity to track demand for future flights and adjust prices accordingly to balance supply and demand.
- Prices are changed by opening and closing predefined booking classes. This is a by-product of the design of the computerized distribution systems, such as SABRE and Galileo, that the airlines developed. These systems allow airlines to establish

a set of prices (fare classes) for each flight and then open or close those fare classes as they wish. This is somewhat different from the pricing issue in most industries, which is not “What fares should we open and close?” but “What prices should we be offering now for each of our products to each market segment through each channel?” The difference is subtle, but it leads to major differences in system design and implementation.

Many companies are understandably wary about adopting “revenue management” programs, protesting that “we are not an airline.” In general, this is the right view—the algorithms behind airline revenue management do not transfer directly to most other industries. However, the experience of the airlines contains several important lessons.

- Pricing and revenue optimization can deliver more than short-term profitability benefits. Revenue management enabled American Airlines to meet the challenge posed by PeopleExpress. It also meant the difference between survival and bankruptcy for National Rent-a-Car. In 1992, National was losing \$1 million per month and was on the verge of being liquidated by its then-owner, General Motors.⁹ At this point, National had been through two rounds of downsizing, and corporate management felt there were no more significant savings that could be achieved on the cost side. As a last-ditch effort, National decided to work on the revenue side. They worked with the revenue management company Aeronomics to develop a system that forecast supply and demand for each car type/rental length at all 170 corporate locations and adjusted fares to balance supply and demand. The results were immediate.

National initiated a comprehensive revenue management program whose core is a suite of analytic models developed to manage capacity, pricing, and reservation. As it improved management of these functions, National dramatically increased its revenue. The initial implementation in July 1993 produced immediate results and returned National Car Rental to profitability. (Geraghty and Johnson 1997)

- E-commerce both necessitates and enables pricing and revenue optimization. The airlines pioneered electronic distribution—their computerized distribution systems, SABRE and Galileo, were the “Internet before the Internet.” These systems allowed immediate receipt and processing of customer booking requests. They also enabled airlines to change prices and availability and have the updated information instantaneously transmitted worldwide. In effect, the airlines were wrestling with the complexities of e-commerce well before the arrival of the Internet. The necessity to continually monitor demand and update prices accordingly will be felt by more and more industries as electronic distribution channels such as the Internet become more pervasive.
- Effective segmentation is critical. The key to the success of revenue management in the airline industry was the ability of the airlines to segment customers between early-booking leisure passengers and late-booking business passengers. Note that this segmentation was achieved not by direct discrimination—that is, trying to charge

a different fare based on demographics, age, or other customer characteristics—but via product differentiation, creating different products that appealed to different segments. Segmenting customers based on their willingness to pay and finding ways to charge different prices to different segments is a critical piece of pricing and revenue optimization—one that we address in detail in Chapter 4.

At heart, airline yield management systems are highly sophisticated opportunity cost calculators. They forecast the future opportunities to sell a seat and seek to ensure that the seat is not sold for less than the expected value of those future opportunities. Most industries do not face capacity constraints as stark as those faced by the airlines. Manufacturers typically have the opportunity to adjust production levels or store either finished or partially finished goods. Retailers can adjust their stocks in response to changes in demand. However, this does not mean that calculating opportunity cost is not relevant to these industries. On the contrary, in many industries facing inventory or capacity constraints, opportunity cost can be the critical link between supply chain management and pricing and revenue optimization. We explore this link in greater detail in Chapter 5.

1.1.2 The New Wealth of Information

As airlines began to develop more sophisticated revenue management systems in the 1990s, other businesses were adopting a new generation of corporate software. Historically, most business software had been homegrown, highly specialized, and oriented toward a single corporate function, such as payroll or invoicing. Such “legacy systems” had often been developed in isolation, were not integrated, and used independent and often-conflicting data sources and definitions. As a result, many companies were reaching the point of diminishing returns on their information technology (IT) investments, with an increasing amount of each IT dollar spent on integrating existing systems rather than building new functionality. Furthermore, the lack of consistent information among departments was often leading to inefficiency and frustration, with identical data entered several times and data items with the same name often having different meanings or interpretation.

This set the stage for enterprise resource planning (ERP) systems. Vendors such as Oracle, SAP, PeopleSoft and Baan began offering integrated client-server ERP systems that enabled corporations to obtain a unified view of all their data. These systems enabled different parts of a company to have access to a consistent, definitive view of corporate data. This in turn enabled efficient and consistent cross-functional business processes without the need for different groups to rekey data or access disparate software systems. The ERP vision is to provide a corporate “information backbone” that supports all business users with consistent and timely data from a single source. Information will then flow freely among business processes. Analytical applications such as supply chain management or staffing systems sit on top of the ERP system and draw the needed inputs from the ERP database. Information “islands” and “fiefdoms” are eliminated, as is the tedious task of shuttling between different software systems and data sources. Corporations will be more efficient, more nimble, and more customer responsive, and IT development can focus on improving functionality rather than integration.

The ERP vision is compelling, but the history of ERP systems has hardly been one of unqualified triumph—many companies have found that replacing their legacy systems with an ERP system can be expensive, disruptive, and painful. Nonetheless, it is clear that the last decade has seen a much-needed consolidation of corporate information. This makes it much easier to implement analytical systems such as pricing and revenue optimization. Such PRO systems require both timely information about product costs and availabilities and the results of recent transactions. By automating and standardizing data consolidation and reconciliation, ERP systems enable much more rapid implementation of pricing and revenue optimization.

Another source of improved data storage and availability came from the increasing adoption of *customer relationship management* (CRM) systems. These involve gathering and storing customer and transaction information and using the results to improve marketing, sales, and customer service. Customer relationship management systems from such vendors as Siebel and e-piphany collect customer and transaction data from different channels and make it available in a data warehouse to various business intelligence, data-mining, and analytic systems.

Pricing and revenue optimization systems are a natural extension of CRM. In essence, CRM systems provide the rich customer and transaction history that pricing and revenue optimization systems need to evaluate customer response and update pricing recommendations. Harrah's Entertainment, which operates 24 casinos and 16 hotels under the Harrah's, Harvey's, Rio, and Showboat brands, provides an example of the successful use of a CRM system with pricing optimization. Harrah's has linked a homegrown CRM system with its reservation and revenue management systems. Based on historical information about customer behavior and preferences, customers are classified into 64 segments based not only on their current value but on their expected lifetime value to Harrah's. The revenue management system forecasts daily hotel room demand for each of these segments and calculates the minimum room price necessary to optimize the room inventory. When customers call for reservations, call-center representatives can see on their screens the customer segment and the approved pricing offers.

The Harrah's system provides an excellent example of a CRM system closely linked with a price optimization system. The CRM system tracks customer behavior and demographics information, such as zip codes, which enables segmentation of customers into such groupings as "avid experienced player" and the calculation of expected profitability and gaming revenue from each segment. The revenue management system balances the need to ensure that rooms will be available for high-revenue customers while not holding back so many rooms that occupancy suffers. The information needed to manage customers at this level of detail simply would not have been available before the advent of customer relationship management technology.¹⁰ Harrah's is an excellent example of how the new wealth of customer information available through CRM systems can enable pricing and revenue optimization.

1.1.3 The Rise of e-Commerce

By the late 1990s, the Internet was widely predicted to be a "revolutionary" and transformative technology that would "change everything." The fact that the Internet will drive a

TABLE 1.3
Results of two studies on the average impact of a 1%
improvement in different variables on operating profit

	McKinsey (1992)	A. T. Kearney (2000)
Price management	11.1%	8.2%
Variable cost	7.8%	5.1%
Sales volume	3.3%	3.0%
Fixed cost	2.3%	2.0%

The passenger airlines typically claim between 8% and 11% benefit from the use of their revenue management systems (Smith et al. 1992)—a figure remarkably consistent with the numbers shown in Table 1.3. A *Harvard Business Review* article noted that some retailers are achieving “gains in gross margins in the range of 5–15%” from the use of optimization-based assortment and pricing optimization systems. Early adopters include J. C. Penney and Gymboree (Friend and Walker 2001). ShopKo Stores has seen a 24% increase in gross margin from the pilot use of a markdown management system (Levison 2002). Ford Motor Company has used a promotions management system to significantly improve the use of its \$9 billion annual North American promotions budget (Leibs 2000) and expects to realize an additional \$5 billion in profits over 5 years from effective market segmentation.¹⁷

Putting it all together, we can see there is a strong case for many companies to consider pricing and revenue optimization. Not only is improving pricing already the “fastest and most cost-effective way to increase profits,” but it is likely to gain in importance as the velocity and complexity of pricing decisions inexorably increase. Furthermore, a new generation of information technology provides the information and algorithmic power needed to analyze and exploit market opportunities. Finally, not only does pricing have extremely high leverage in improving profitability, as shown in Table 1.3; it is often the area that can be improved the most with the least investment. The focus on cost improvement throughout the 1990s means that incremental improvements in variable costs and fixed costs will be more expensive and difficult than incremental improvements in pricing and price management.

1.3 ORGANIZATION OF THE BOOK

The structure and dependencies of the remaining chapters are illustrated in Figure 1.1. The topic of each chapter is outlined briefly next.

Chapter 2 discusses pricing and revenue optimization as a corporate process and contrasts it with other approaches to pricing. It stresses that PRO is a highly dynamic process, dependent on continual feedback to ensure that pricing decisions are kept in line with changing market realities. It also discusses the concept that the core of pricing and revenue optimization lies in the formulation of pricing and availability decisions as constrained optimization problems.

Chapter 3 reviews the basic economics behind price optimization. It introduces the idea of a price-response curve and presents several common forms of price-response curve. It shows how the price-response curve can incorporate competitive pricing. It argues that *incremental cost* is also a critical input to pricing a customer commitment. Finally it shows

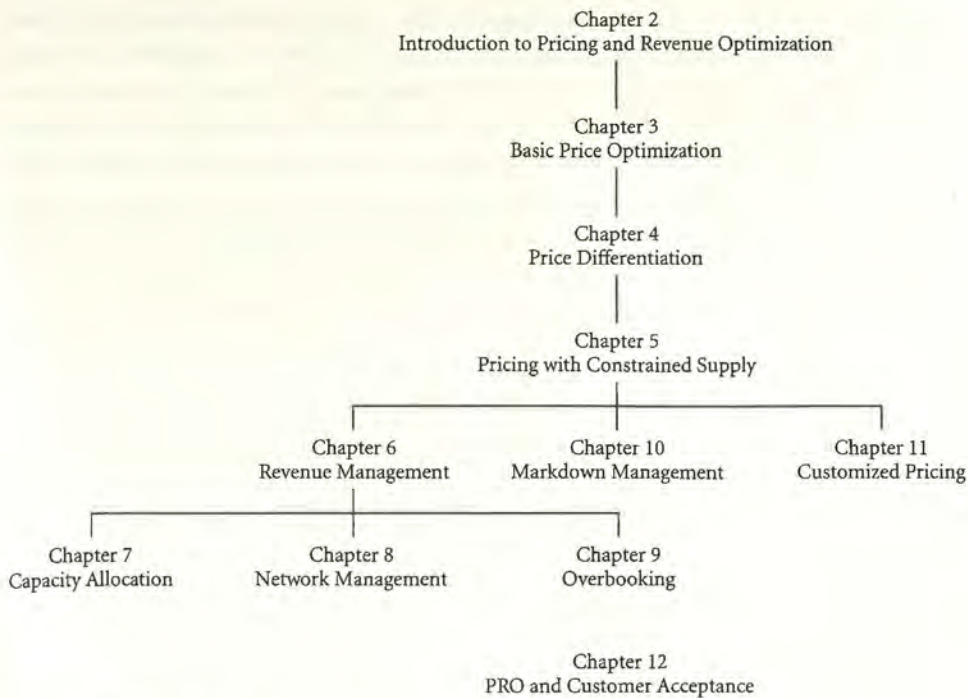


Figure 1.1 Relationships among the remaining chapters.

several ways that prices can be optimized in the simplest case of a supplier selling a single product with a known price-response curve.

Price differentiation is at the core of pricing and revenue optimization. Chapter 4 discusses how markets can be divided into different market segments such that a different price can be charged to each segment. Tactics for price differentiation include virtual products, product lines, group pricing, channel pricing, and regional pricing. The idea of *multipricing* is the source of both the benefits of PRO and the challenges involved in managing a large portfolio of prices. This chapter discusses how to optimize differentiated prices in the face of potential cannibalization.

Chapter 5 introduces another major theme in pricing and revenue optimization—pricing when supply (or inventory) is constrained. Supply constraints are *ubiquitous*. They may arise from the intrinsically constrained capacity of a service provider, such as a hotel, restaurant, barbershop, or trucking company; they may be due to limited inventory on hand; or they may be due to “bottlenecks” in a supply chain that restrict the rate at which a good can be produced or transported. In any case, constrained supply significantly complicates optimal pricing. This chapter also introduces the key concept of an *opportunity cost*—the incremental contribution lost due to a supply constraint.

Revenue management has been one of the most important and publicized applications of pricing and revenue optimization over the past two decades. While it has its origin in the airline industry, it has spread beyond the airlines and is in widespread use at hotels, rental car firms, cruise lines, freight transportation companies, and event ticketing agents. Some

of the core techniques are gaining acceptance and use well beyond these industries. We devote four chapters to revenue management. The first—Chapter 6—describes the background, history, and business setting of revenue management. Chapter 7 is devoted to capacity allocation, the techniques used to determine which fare classes should be open and which closed at any time for a product consisting of a constrained resource. Chapter 8 extends this analysis to the case of a network, in which an individual product may use many different resources. Chapter 9 discusses overbooking—the question of how many units of a constrained product should be sold when customers may not show or may cancel.

Markdown management—the topic of Chapter 10—is an increasingly popular application of pricing and revenue optimization. In a markdown industry, a merchant has a stock of inventory whose value decreases over time. His problem is to determine the schedule of price reductions to take in order to maximize the return from inventory. Applications are widespread, from fashion goods, through consumer electronics and durables, to theater tickets. This chapter describes basic markdown optimization models and some of the challenges in implementing markdown management in the real world.

Chapter 11 treats customized prices. Customized pricing is common in business-to-business settings where goods and services are sold based either on long-term contracts or as part of large individual transactions. In these settings, list prices may not exist or may serve only as “guidelines” off of which discounts are set. The pricing and revenue optimization challenge is to determine the discount levels to provide to each customer in order to maximize the *expected profitability* of the deal. This requires estimating the tradeoff between the probability of winning the deal and the margin contribution if the deal is won.

Chapter 12 treats the issue of customer perception and acceptance of pricing tactics. Much of the classical theory of pricing is based on the idea that consumers are rational “utility maximizers.” Prices are emotionally neutral signals that guide purchasing decisions. However, both common sense and recent research has shown that this view is incomplete at best and misguided at worst. Consumers can care deeply about prices and the way they are presented. In particular, pricing that is perceived as “unfair” can trigger an emotional rejection. Chapter 12 discusses the implication of consumer “irrationality” for PRO.

under Sales. When they wanted to increase profits, they would move it into the Finance Department.⁷ Other companies are less disciplined—their approach to pricing may change with the “flavor of the month”: market-based when the emphasis is on market share, value-based when “focusing on the customer” comes into vogue.

Vacillating among pricing approaches is actually better than strict devotion to one approach. Any company that sticks tenaciously to any one of the three pure approaches would likely find itself in deep trouble very quickly. What is often seen in reality is a hybrid—companies espouse a particular philosophy but use pieces of all three, supplemented by a considerable amount of improvisation. The upshot is pricing confusion—there is rarely a consistent justification or approach applied across all pricing decisions.

2.3 THE SCOPE OF PRICING AND REVENUE OPTIMIZATION

Pricing and revenue optimization provides a consistent approach to pricing decisions across the organization. This means that a company needs to have a clear view of all the prices it is setting in the marketplace and the ways in which those prices are set. This defines the scope of pricing and revenue optimization.

2.3.1 The PRO Cube

The goal of pricing and revenue optimization is to provide the right price,

- for every product
- to every customer segment
- through every channel

and to update those prices over time in response to changing market conditions. The three dimensions of pricing and revenue optimization can be illustrated in a cube, as shown in Figure 2.4. Each element within the cube represents a combination of product, channel, and customer segment. Each element (or *cell*) has an associated price. For example, one element might be:

medium-size turbines sold to large customers in the northeast via the direct sales channel

Another element might be:

Replacement gears sold to small companies via online sales

In theory, each cell within the PRO cube could correspond to a different price. In practice, some cells may not be meaningful. Some products may not be offered through some channels, for example. It is also the case, of course, that the prices within the PRO cube will not always be independent of each other. Our ability (or desire) to charge different prices through different channels may be constrained either by practical considerations or by strategic goals. If we want to encourage small customers to purchase online rather than through our direct sales channel, we might institute a constraint that says that the online price for small customers for all products must be less than or equal to the direct sales chan-

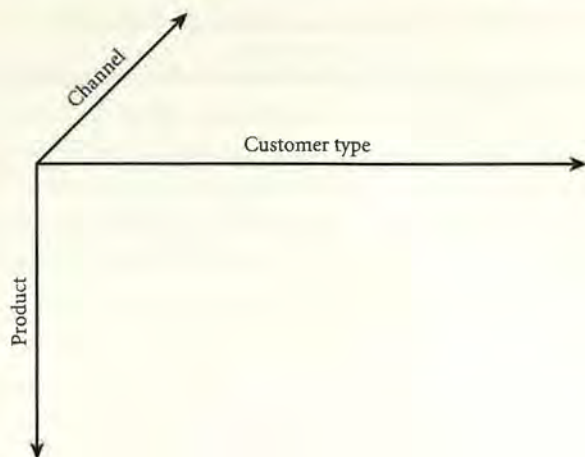


Figure 2.4 · Dimensions of the pricing and revenue optimization cube.

nel price. These considerations are among the business rules that must be considered in the pricing and revenue optimization process described in Section 2.4.

Companies may offer even more prices than the PRO cube would imply. Certain products might be subject to tiered pricing or volume discounts. Or a company might offer bundles of products at different promotional rates or discounts that are not available for individual products. Each of these bundles or quantity combinations can be treated as an additional “virtual product” in the PRO cube.

The PRO cube is a useful starting point for a company seeking to understand the magnitude of the pricing challenge that it faces. Enumerating the combinations of products, market segments, and channels gives a rough estimate of the total number of prices a company needs to manage. Many companies are astonished at the sheer volume of prices they already offer. This astonishment is often the first step in realizing the need to establish a consistent pricing and revenue optimization process.

2.3.2 Customer Commitments

A core concept in pricing and revenue optimization is the idea of a *customer commitment*. Specifically, a *customer commitment* occurs whenever a seller agrees to provide a customer with products or services, now or in the future, at a price. The elements of a customer commitment include:

- The products and services being offered
- The price
- The time period over which the commitment will be delivered
- The time for which the offered commitment is valid—i.e., how long the customer has to make up his mind
- Other elements of the contract or transaction (e.g., payment terms, return policy)
- Firmness of the commitment and risk sharing

Some examples of customer commitments include the following:

- A list price is possibly the most familiar form of customer commitment. A list price is a commitment that a buyer can obtain the item simply by paying the posted price. List price is often (but not always) a “take it or leave it” or “nonnegotiable” price. Common examples are the shelf prices at the grocery store and drugstore, price per gallon displayed on the gasoline pump, and many online retail prices.
- Coupons and other types of promotion are also forms of customer commitment. They allow certain customers to obtain a good at a price lower than list for some period.
- In a business-to-business setting, prices for large purchases are often individually negotiated. The seller may have published list prices, but, in many industries, items are rarely, if ever, purchased at the list price. Rather, each buyer negotiates an individual discount. The discount that a particular buyer receives can depend on the buyer’s purchasing history with the seller, the skills of the purchasing agent and the salesperson, and the desire of the seller to make the sale. We describe an approach to optimizing such customized prices in Chapter 11.
- In other business-to-business settings, the buyer and seller negotiate a contract that establishes prices that will be in place for six months, a year, or longer. For example, in less-than-truckload (LTL) freight, shippers agree on a schedule of tariffs that will govern all the shipments they will make over the next year. These schedules are usually expressed in terms of a discount from a standard list tariff. Contracts are individually negotiated between each shipper and each carrier. Contracts may be exclusive (as they often are in package express), or a shipper may establish contracts with two or more carriers, as is common practice in trucking and container shipping.
- Contracts between electronics distributors, such as Ingram Micro and TechData, and wholesalers, such as Hewlett-Packard and Sun Microsystems, are often on a “tiered pricing” basis, where each tier includes a volume target and an incremental discount for hitting that target. An example is shown in Table 2.2. In this case, the purchaser will receive at least a 5% discount on each unit purchased during the current quarter. If she purchases at least 50,001 units, the discount increases to 6%. And if she purchases 80,001 units, the discount increases to 7%. This type of tiered pricing provides an incentive for the distributor to “push” the seller’s products rather than those of its competitors. Tiered discounts can be based on either units sold or revenue sold. The increased discounts themselves can apply either to incremental sales above the tier or to all sales during the quarter.
- Online auction houses such as eBay enable a seller to establish a reserve price for an item they wish to sell. On eBay, the highest bidder above the reserve price will be sold the item at the maximum of the second highest bidder’s price and the reserve price. The commitment by the seller is to honor this policy. In addition, the seller may establish a “buy now” price that allows immediate purchase. Auctions, either online or offline, are also common for many types of industrial goods.

TABLE 2.2
Tiered pricing

Quarterly units sold	Total discount (%)
0–50,000	5
50,001–80,000	6
80,001–110,000	7
≥110,001	8

The forms and types of commitments that sellers make (and buyers expect) vary from industry to industry. As a simple example, airline tickets were historically refundable, meaning that the airline took all the risk of a customer “no-show”—a risk the airlines sought to mitigate via overbooking. On the other hand, tickets to Broadway shows have historically been nonrefundable, meaning the customer takes on the risk of no-show. As a result, Broadway theaters do not overbook. The details of the types of commitments made in different industries are often the result of complex and contingent historical factors and may change over time—just as airlines are increasingly beginning to adopt nonrefundable ticketing.

In fact, it is not uncommon for individual sellers to be utilizing different approaches for different products through different channels. For example, an automobile manufacturer is likely to sell the majority of vehicles through dealers, utilizing a combination of “shelf price” (in this case, the wholesale price) and promotions. However, Ford Motor Company receives about 9% of its revenue in North America from group sales of cars to corporate, government, and industrial fleets. These sales are generally priced through individual negotiations or bids. Finally, a manufacturer may also be selling some cars directly to the public through a deal with an online channel. In this case, the manufacturer needs to apply pricing and revenue optimization across the entire range of its channels and pricing mechanisms in order to maximize total profitability.

2.4 THE PRICING AND REVENUE OPTIMIZATION PROCESS

Successful pricing and revenue optimization involves two components:

1. A consistent business process focused on pricing as a critical set of decisions
2. The software and analytical capabilities required to support the process

Much of the interest in pricing and revenue optimization has focused on its use of mathematical analysis. Quantitative analysis is indeed critical in most pricing and revenue optimization settings, but it cannot provide sustainable improvement unless it is embedded in the right process. An overview of such a process is shown in Figure 2.5. In this picture, the overall pricing and revenue optimization process has been divided into eight activities. Four of these activities are part of “operational PRO”—that is, they are executed every time the company needs to change the prices it is offering in the marketplace. The remaining four “supporting PRO” activities occur at longer intervals. As the figure shows, pricing and revenue optimization is a closed-loop process; that is, feedback from the market must be

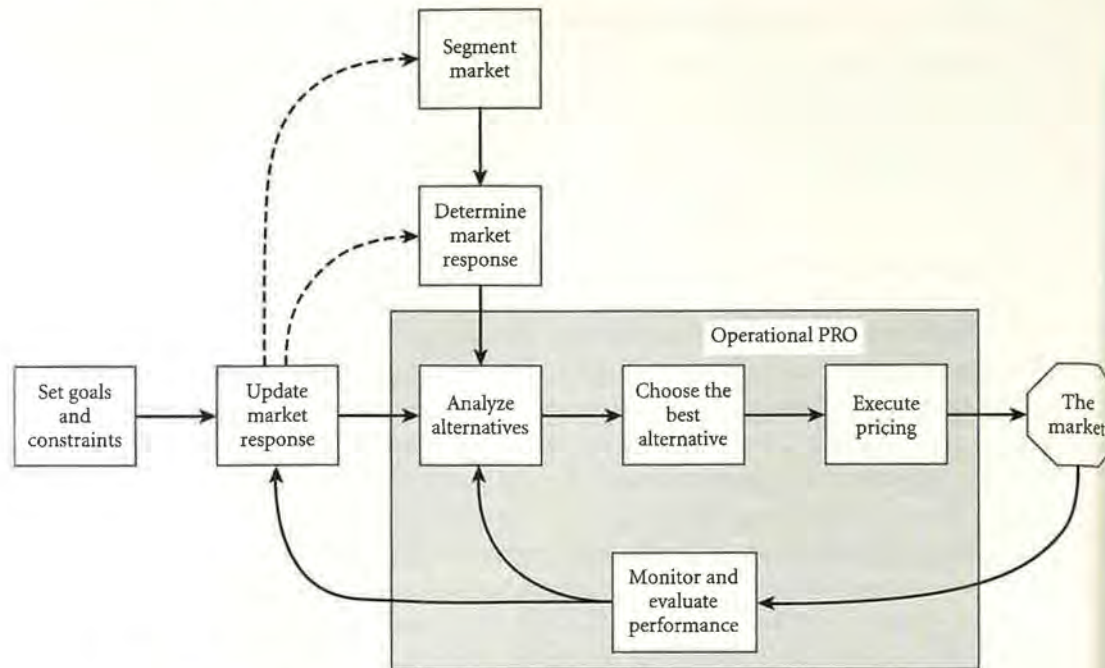


Figure 2.5 The pricing and revenue optimization process.

incorporated into both the operational activities and the more periodic activity of updating market response curves.

2.4.1 Operational PRO Activities

The operational PRO activities work continuously to set and update prices in the marketplace. The timing of the decisions will depend on the application—airline revenue management systems can change fares from moment to moment. Revenue managers review the recommendations of the system on a daily basis. Other companies may update prices weekly or monthly. In any case, the core operational activities will be similar.

Analyze alternatives. This activity is the one most widely identified with pricing and revenue optimization or revenue management. Historically it has involved analysts' use of spreadsheets to compare pricing alternatives under different scenarios. More recently it has involved the use of software systems that solve the underlying optimization problem to recommend new prices for every element of the PRO cube.

Choose the best alternative. However prices are analyzed, a human being needs to be responsible for determining which prices should be set. For this reason, almost all pricing and revenue optimization software packages are equipped with a "what-if" capability that allows analysts to understand the reasons for their recommendations and the effect of different assumptions on the recommendations. In very rapidly changing environments such as the passenger airlines, the majority of prices will be updated automatically, with analysts only investigating the unusual or unexpected situations.

Execute pricing. Finally, the prices that have been calculated need to be communicated to the market. The mechanics of price transmission and distribution vary greatly from industry to industry, company to company, and even channel to channel within the same company. For an airline or hotel, prices are updated by opening or closing fare classes—a process discussed in more detail in Chapter 6. In other cases, a “pricing matrix” database specifies which prices are available to which customers through which channels. In other cases, new discount guidelines need to be communicated to the sales force. For a company with many products being sold through many channels, communicating the correct prices correctly can be a challenge in itself. A number of pricing execution or “pricing management” vendors, such as Vendavo and Metreo, sell software whose primary function is to ensure that complex pricing changes are correctly and efficiently transmitted through all channels.

Monitor and evaluate performance. As we receive the results from the marketplace, we need to compare the results with our expectations and evaluate overall performance against our goal. Are we achieving the lift we anticipated from a promotion? If not, why not? Was total market demand less than expected? Did the competition react more aggressively than we had anticipated? Were customers less responsive to price than we anticipated? Or was it a combination of all three?

Figure 2.5 shows that market feedback occurs at two levels. The most immediate feedback is to the analysis of alternatives. Here, the effects of the most recent actions should be monitored so that immediate action can be taken if necessary. The required actions may not always be directly related to pricing and revenue optimization. For example, the markdown management system utilized by a major retailer indicated that a particular style of pants was selling well below expectations. Investigation found that the pants had been hemmed improperly. Once the problem had been corrected, the pants began to sell at the rate initially expected.⁸ In the more common case, a big difference between realized sales and expected sales is likely to be due to an unforeseen change in the marketplace.

The second level of feedback updates the parameters of the market-response functions. If sales of some product are slower than expected, it may indicate that the market is more price responsive than expected and the future market-response curve for that product should be adjusted accordingly.

2.4.2 Supporting PRO activities

The primary role of the supporting activities is to provide key inputs to the operational PRO activities. The supporting activities occur on a much longer time frame. Typically goals and business rules will change quarterly at best. A new market segmentation may take place annually. Price-response curves will be updated much more frequently, typically weekly or monthly in the case of markdown management or customized pricing.

Set goals and business rules. A key initial step in pricing and revenue optimization is to specify the overall goal of the process. Without a clear goal it is impossible to make consistent decisions and it is impossible to evaluate decisions and improve the process over time. In general, the goal of pricing and revenue optimization is to maximize expected total contribution. However, at certain times, in certain markets, a company may wish to increase market share or attain a volume sales goal. Whatever the goal, it is most important that it be

stated clearly and explicitly. As we discuss in Chapter 3, the corporate goal in a product/segment/channel combination determines the form of the objective function that will be used when optimizing that “cell” of the PRO cube.

Business rules apply whenever there is a restriction that may prevent us from pricing “optimally” in some market. Is there a minimum per-unit margin requirement we need to meet for all products within a category? Is there a minimum discount we need to offer to certain strategic customers? Do we want our retail list prices to be no lower than our Internet prices? Each of these types of business rule needs to be implemented as a constraint in the optimization problem at the heart of PRO.

Example 2.1

A heavy-equipment manufacturer wants to make sure that prices in the northeast are never more than 15% higher than the comparable model being offered by the leading competitor. This decision results in a constraint to be included in the PRO optimization problem:

Our price in the northeast $\leq$ 115% competitor A's price in the northeast.

We note that business rules are *inputs* into PRO—that is, they are decisions that are made not on the basis of short-term optimization but on the basis of other considerations a company wishes to include in its pricing.

Segment the market. An important theme within pricing and revenue optimization is segmenting the market in a way that maximizes the opportunity to extract profit. We discuss techniques for doing this in Chapter 4. The market segmentation needs to be updated periodically in order to reflect changes in the underlying market. However, this updating should be performed much less frequently than the operational PRO processes. For this reason, the “Segment Market” step in Figure 2.5 is shown with a dashed line outside the main process loop.

Determine price response. For each market segment identified we need to calculate the corresponding price-response functions. This step is performed in conjunction with segmenting the market.

Update price response. It is not sufficient merely to monitor the performance of PRO actions against expectations. A company needs to update its models to incorporate what they learn. As long as performance closely matches expectations, the parameters of the models need not be changed. However, if performance is significantly different from expectations, models need to be updated in order to reflect the new information.

The most critical message to take away from the pricing and revenue optimization process illustrated in Figure 2.5 is that PRO should be treated like any other business process. To make effective PRO decisions, we need to identify clearly what it is trying to achieve, the constraints we are facing, and the alternatives available. Based on our understanding of the market and the constraints, we choose the alternative most likely to achieve our goals. Once this alterna-

tive is implemented, we must monitor and measure the results against our expectations and update our understanding of the market so that we can make better decisions in the future. The PRO organization, responsibilities, systems, and data-capture methodologies all need to be able to support this process.

The PRO process presented in this section is highly general. We return to it in future chapters as we look at how pricing and revenue optimization is applied in specific pricing situations, such as revenue management, customized pricing, and markdown management.

2.4.3 The Time Dimension

Much of the growing interest in pricing and revenue optimization is being driven by the increasing velocity of pricing decisions. Where, in the past, prices changed once a quarter for many industries, now they change once a week, once a day, or even once an hour. Increased frequency and complexity of pricing within an industry creates its own momentum. A company that can change its prices more rapidly in response to changing market conditions will gain an advantage. This creates a strong incentive for other companies to match (or exceed) the frequency of change. E-commerce has been a proven driver of increased price velocity. As discussed in Chapter 1, the adoption of computerized distribution systems—an early form of e-commerce—by the passenger airlines was a direct contributor to the ability of the airlines to develop, manage, and update highly complex fare structures.⁹

Pricing has a different cadence in every industry. The rack (wholesale) price for gasoline fluctuates randomly, with little obvious trend. Fashion goods are priced high at the beginning of the season and are then subsequently marked down as the season progresses. Different pricing approaches need to be used in each of these markets. Gasoline rack prices are set using “dynamic pricing” approaches that primarily track changes in the underlying supply and demand imbalance. When supply is low (Iraqi oil production is interrupted) or demand is high (a cold winter in the northeast), the price tends to rise. Opposite conditions cause the price to fall. The price of a fashion item falls across the season, since such items become less valuable as the season progresses.¹⁰ On the other hand, the passenger airlines have successfully segmented their customer base into early-booking leisure customers and later-booking business customers. To support this segmentation, ticket prices rise as departure approaches—a phenomenon we look at in much greater detail in Chapter 6. These different cadences lead to very different optimization approaches in different markets.

Frequent price changes creates a challenge for an organization. As price changes become more frequent and the pricing structure becomes more complex, it becomes less and less feasible for a company to apply in-depth analysis to each pricing decision. The “analyst with a spreadsheet” method of pricing decision making begins to break down. As the complexity and rapidity of pricing decisions continues to accelerate, many organizations find themselves overwhelmed. A common symptom is that despite expending more and more effort on pricing decisions, the company continually seems to be one step behind the market and the competition. At this point, most companies consider adopting some sort of computerized pricing support system.

In this section we introduce the basic elements of pricing and revenue optimization and show how the basic pricing revenue and optimization problem can be formulated as an optimization problem. The goal of the optimization problem—its objective function—is to maximize contribution: total revenue minus total incremental cost from sales. The key elements of this problem are the price-response function and the incremental cost of sales, both of which we introduce in this section. We then formulate and solve the pricing and revenue optimization function in the case of a single product in a single market without supply constraints and derive some important optimality conditions.

3.1 THE PRICE-RESPONSE FUNCTION

A fundamental input to any PRO analysis is the *price-response function*, or *price-response curve*, $d(p)$ which specifies how demand for a product varies as a function of its price, p . There is one price-response function associated with each element in the PRO cube—that is, there is a price-response function associated with each combination of product, market-segment, and channel. The price-response function is similar to the market demand function found in economic texts. However, there is a critical difference. The price-response function specifies *demand for the product of a single seller as a function of the price offered by that seller*. This contrasts with the concept of a *market demand curve*, which specifies how an entire market will respond to changing prices. The distinction is critical because different firms competing in the same market face different price-response functions. Referring to Table 1.2, the price-response function facing Amazon for *Bag of Bones* is likely to be quite different from the one facing ecampus.com. The differences in the price-response functions faced by different sellers are the result of many factors, such as the effectiveness of their marketing campaigns, perceived customer differences in quality, product differences, and location, among other factors.

In a perfectly competitive market, the price response faced by an individual seller is a vertical line at the market price, as shown in Figure 3.1. If the seller prices above the market

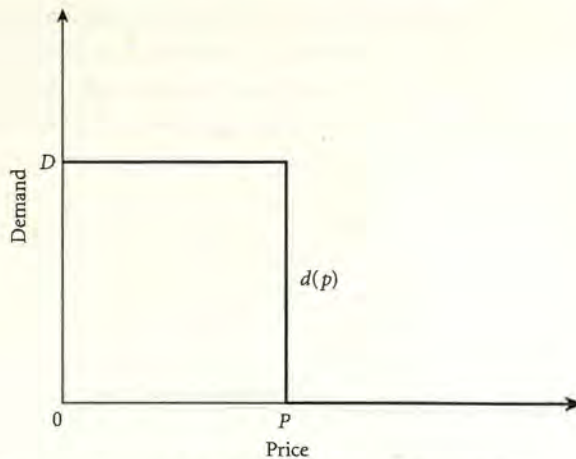


Figure 3.1 Price-response curve in a perfectly competitive market.

price, his demand drops to 0. If he prices below the market price, his demand is equal to the entire market. A standard example used in economics texts is wheat:

The best example to keep in mind is that of a wheat farmer, who provides a minuscule percentage of the wheat grown in the world. Regardless of whether he produces 10 bushels or 1,000, he remains too small to have any impact on the going market price. . . . If he tries to charge even a fraction of a penny more, he will sell no wheat, because buyers can just as easily buy from someone else. If he charges even a fraction of a penny less, the public will demand more wheat from him than he can possibly produce—effectively an infinite quantity.¹

In other words, wheat is a commodity—buyers are totally indifferent among the offerings of different sellers, they have perfect knowledge about all prices being offered, and they will buy the product only from the lowest-price seller. Furthermore, each seller is small relative to the total size of the market. In this situation, the seller has no pricing decision—his price is set by the operation of the larger market. To quote a popular text: “In a competitive market, each firm only has to worry about how much output it wants to produce. Whatever it produces can only be sold at one price: the going market price.”² At any price below the market price, the demand seen by a seller would be equal to the entire demand in the market—the amount D in Figure 3.1. At any price above the market price he sells nothing. The seller of a true commodity in a perfectly competitive market has no need for pricing and revenue optimization—indeed, he has no need of any pricing capability whatsoever. However, true commodities are surprisingly rare. The vast majority of companies face finite customer responses to price changes and therefore have active PRO decisions.

The price-response functions that we consider—and those facing most companies most of the time—demonstrate some degree of smooth price response. An example is shown in Figure 3.2. Here, as price increases, demand declines until it reaches zero at some *satiating price* P . This type of smooth market-response function is usually termed a *monopolistic* or *monopoly* demand curve in the economics literature. The terminology is somewhat unfortunate, since the fact that a company faces some level of price response hardly means that

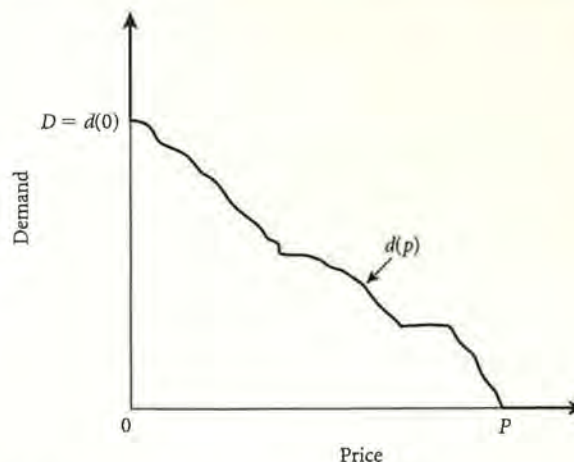


Figure 3.2 Typical price-response curve.

it is a “monopoly.” Companies such as United Airlines, UPS, and Ford all face smooth price-response functions for their products, yet none of them would be considered a monopoly in the usual sense of the world.

3.1.1 Properties of the Price-Response Function

The price-response functions used in PRO analysis have a time dimension associated with them. This is in keeping with the dynamic nature of PRO decisions—we are not setting prices that will last “in perpetuity” but prices that will be in place for some finite period of time. The period might be minutes or hours (as in the case of a fast-moving e-commerce market), days or weeks (as in retail markets), or longer (as in long-term contract pricing). At the end of the period we have the opportunity to change prices. The demand we expect to see at a given price will depend on the length of time the price will be in place. Thus, we can speak of the price-response function for a model copier over a week or over a month, but without an associated time interval there is no single price-response function.

There are many different ways in which product demand might change in response to changing prices and, thus, many different possible price-response functions. However, all of the price-response functions we consider will be nonnegative, continuous, differentiable, and downward sloping.

Nonnegative. We assume that all prices are greater than or equal to zero; that is, $p \geq 0$.³

Continuous. We assume that the price-response function is *continuous*—there are no “gaps” or “jumps” in the market response to our prices. More formally, if $d(0) = D$ and P is the satiating price, that is, the lowest price for which $d(P) = 0$, then for every $0 < q \leq D$ there is a price $p \leq P$ such that $d(p) = q$. This implies that there is a price that will generate any level of demand between 0 and the maximum demand. This property, known as *invertability*, is often very useful.⁴

Differentiable. Differentiability means that the price-response function is smooth, with a well-defined slope at every point.⁵ As with the assumption of continuity, this assumption involves taking a mathematical liberty, since prices are only defined for fixed increments.

However, differentiability allows us to use the tools of calculus to solve the constrained optimization problems that arise in PRO—a gain that outweighs the slight imprecision that results from using derivatives rather than difference equations.

Downward sloping. We assume that $d(p)$ is downward sloping whenever $d(p) > 0$ —that is, that raising prices for a good during a period will decrease demand for that good during that period unless demand is already 0, in which case demand will remain at 0. Conversely, lowering prices can only increase demand.

The “downward sloping” assumption calls for a bit more discussion. First of all, it should be noted that downward-sloping demand curves do not mean that high prices will always be associated with low demand. A hotel revenue management will experience higher average rates when occupancy is high and lower average rates when occupancy is low. What the downward-sloping property does indicate is that, in any time period, demand would have been lower if prices had been higher, and vice versa. This corresponds both to economic theory (in which consumers maximize their utilities subject to a budget constraint) and to real-life experience.

Nonetheless, there are at least three cases in which the downward-sloping property may not hold.

1. *Giffen goods.* Economic theory allows for the possibility of so-called *Giffen goods*, whose demand rises as their price rises because of substitution effects. An example might be a student on a strict budget of \$8.00 per week for dinner. When hamburger costs \$1.00 per serving and steak costs \$2.00, he eats hamburger six times a week and steak once. If the price of hamburger rises to \$1.10, to stay within his budget, he stops buying steak and buys hamburger seven times a week. In this case, a rise in the price of hamburger causes his consumption of hamburger to increase. While this behavior might conceivably occur at an individual level, a Giffen good requires that enough buyers act this way that they overwhelm other buyers who would buy less hamburger as the price rises. Giffen goods are almost never encountered in reality—in fact, many economists doubt whether they have ever existed.
2. *Price as an indicator of quality.* In some markets, price is used by some consumers as an indicator of quality: Higher prices signal higher quality. In this case, lowering the price for a product may lead consumers to believe that it is of lower quality, and demand could drop as a result. Typically, markets where this is an issue have a large number of alternatives and some “lazy” buyers who do not have the time or resources to research the relative quality of all the alternatives so that they use price as a proxy. Wine is a classic example: Faced with a daunting array of labels and varietals, many purchasers are likely to use a rule such as: a \$10 bottle for dinner with the family, a \$15 bottle if the couple next door is dropping by, and a \$25 bottle if our wine-snob friends are joining us for dinner.⁶ The “price-as-an-indicator-of-quality effect” can be particularly important when a new product enters the market. A medical-product company developed a way of producing a home testing device at a cost 75% below the cost of the prevailing technology. They introduced the new product at a list price 60% lower than the list price of the leading competitors, expecting to dominate the market. When sales were slow, they repackaged the product

and sold it at a price only 20% lower than the leading competitor. This time sales took off. Their belief is that the initial rock-bottom price induced customers to believe their product was inferior and unreliable. The higher price was high enough not to raise quality concerns but low enough to drive high sales.

3. *Conspicuous consumption.* Thorstein Veblen coined the term *conspicuous consumption* for the situation in which a consumer makes a purchase decision in order to advertise his ability to spend large amounts. It probably does not come as a shock that the reason some rock stars drink \$300 bottles of Krystal champagne and drive Bentleys is not their finely honed appreciation of fine French champagne and British automotive engineering. Conspicuous consumption postulates a segment of customers who buy a product simply because it has a high price—and others know it. Dropping the price in this case may cause the product to lose its cachet and decrease demand.

While duly noting these three exceptions to downward-sloping demand curves, we will proceed to ignore them for the remainder of the book. In defense of this decision, we observe that for almost all items, almost all of the time, raising the price will lower demand and lowering the price will increase demand.

3.1.2 Measures of Price Sensitivity

It is often useful to have a simple characterization of the price sensitivity implied by a price-response function at a particular price. The two most common measures of price sensitivity are the *slope* and the *elasticity* of the price-response function.

Slope. The *slope* of the price-response function measures how demand changes in response to a price change. It is equal to the change in demand divided by the difference in prices, or

$$\delta(p_2, p_1) = [d(p_2) - d(p_1)] / (p_2 - p_1) \quad (3.1)$$

By the downward-sloping property, $p_1 > p_2$ implies that $d(p_1) \leq d(p_2)$. This means that $\delta(p_1, p_2)$ will always be less than or equal to zero.

The definition in Equation 3.1 requires two prices to be specified, because the slope of a price-response function will be constant across all prices only if it is linear. However, it is common to specify the slope at a single price, say, p_1 , in which case it can be computed as the limit of Equation 3.1 as p_2 approaches p_1 . That is,

$$\begin{aligned} \delta(p_1) &= \lim_{h \rightarrow 0} [d(p_1 + h) - d(p_1)] / h \\ &= d'(p_1) \end{aligned}$$

where $d'(p_1)$ denotes the derivative of the price-response function at p_1 . By the differentiability property, we know that this derivative exists. The downward-sloping property means that the slope will be less than or equal to zero for all prices.

The slope can be used as a *local* estimator of the change in demand that would result from a small change in price. For small changes in price, we can write

$$d(p_2) - d(p_1) \approx \delta(p_1)(p_2 - p_1) \quad (3.2)$$

That is, a large (highly negative) slope means that demand is more responsive to price than a smaller (less negative) slope.

Example 3.1

The slope of the price-response function facing a semiconductor manufacturer at the current price of \$0.13 per chip is $-1,000$ chips/week per cent. From Equation 3.2, he would estimate that a 2-cent increase in price would result in a reduction in demand of about 2,000 chips per week and a 3-cent decrease in price would result in approximately 3,000 chips/week in additional demand.

It is important to recognize that the quality of the approximation in Equation 3.2 declines for larger changes in prices and that the slope cannot be used as an accurate predictor of demand at prices far from the current price. It is also important to realize that the slope of the price-response function depends on the units of measurement being used for both price and demand.

Example 3.2

The price of a bulk chemical can be quoted in either cents per pound or dollars per ton. Assume that the demand for the chemical is 50,000 pounds at 10 cents per pound but drops to 40,000 pounds at 11 cents per pound. The slope of the price response function at these two points is

$$\begin{aligned}\delta(10, 11) &= (50,000 - 40,000)/(10 - 11) \\ &= -10,000 \text{ pounds/cent}\end{aligned}$$

The same slope in tons per dollar would be $(25 - 20)/(0.1 - 0.11) = -500$ tons/dollar.

Price elasticity. Perhaps the most common measure of the sensitivity of demand to price is *price elasticity*, defined as the ratio of the percentage change in demand to the percentage change in price.⁷ Formally, we can write

$$\epsilon(p_1, p_2) = -\frac{100\{[d(p_2) - d(p_1)]/d(p_1)\}}{100\{(p_2 - p_1)/p_1\}} \quad (3.3)$$

where $\epsilon(p_1, p_2)$ is the elasticity of a price change from p_1 to p_2 . The numerator in Equation 3.3 is the percentage change in demand, and the denominator is the percentage change in price. Reducing terms gives

$$\epsilon(p_1, p_2) = -\frac{[d(p_2) - d(p_1)]p_1}{[p_2 - p_1]d(p_1)} \quad (3.4)$$

The downward-sloping property guarantees that demand always changes in the opposite direction from price. Thus, the minus sign on the right-hand side of Equation 3.4 guarantees that $\epsilon(p_1, p_2) \geq 0$. An elasticity of 1.2 means that a 10% increase in price would result in

a 12% decrease in demand and an elasticity of 0.8 means that a 10% decrease in price would result in an 8% increase in demand.

$\epsilon(p_1, p_2)$ as defined in Equation 3.4 is sometimes called the *arc elasticity*. That it requires two prices to calculate reflects that the percentage change in demand resulting from changing prices will depend on both the old price and the new price. In fact, the percentage decrease in from a 1% increase in price will generally not even be the same as the percentage increase in demand that we would experience from a 1% decrease in price. For this reason, both prices need to be specified in order to fully characterize elasticity. However, as we did with slope, we can derive a *point elasticity* at p by taking the limit of Equation 3.4 as p_2 approaches p_1 :

$$\epsilon(p_1) = -d'(p_1)p_1/d(p_1) \quad (3.5)$$

In words, the point elasticity is equal to -1 times the slope of the demand curve times the price, divided by demand. Since $d'(p) \leq 0$, the point elasticity $\epsilon(p)$ calculated by Equation 3.5 will be greater than or equal to zero. The point elasticity is useful as a local estimate of the change in demand resulting from a *small* change in price.

Example 3.3

A semiconductor manufacturer is selling 10,000 chips per month at \$0.13 per chip. He believes that the price elasticity for his chips is 1.5. Thus, a 15% increase in price from \$0.13 to \$0.15 per chip would lead to a decrease in demand of about $1.5 \times 15\% = 22.5\%$, or from 10,000 to about 7,750 chips per month.

One of the appealing properties of elasticity is that, unlike slope, its value is independent of the units being used. Thus, the elasticity of electricity is the same whether the quantity electricity is measured in kilowatts or megawatts and whether the units are dollars or euros.

Example 3.4

Consider the bulk chemical whose price-response slope was estimated in Example 3.2. It showed a 20% decrease in demand (from 50,000 pounds to 40,000 pounds) from a 10% increase in price (from 10 cents to 11 cents). The corresponding elasticity is $0.2/0.1 = 2$ —an elastic response. What if the units were euros and tons? It would still be a 20% decrease in demand (from 25 tons to 20 tons) from a 10% increase in price.

Like slope, point elasticity is a *local* property of the price-response function. That is, elasticity can be specified between two different prices by Equation 3.4 and a point elasticity can be defined by Equation 3.5. However, the term *price elasticity* is often used more broadly and somewhat loosely. Thus, statements such as “gasoline has a price elasticity of 1.22” are imprecise unless they specify both the time period of application and the reference price. In practice, the term *price elasticity* is often used simply as a synonym for *price sensitivity*. Items with “high price elasticity” have demand that is very sensitive to price while “low price

elasticity" items have much lower sensitivity. Often, a good with a price elasticity greater than 1 is described as elastic, while one with an elasticity less than 1 is described as inelastic.

Elasticity depends on the time period under consideration, and, as with other aspects of price response, we must be clear to specify the time frame we are talking about. For most products, *short-run elasticity* is lower than *long-run elasticity*. The reason is that buyers have more flexibility to adjust to higher prices in the long run. For example, the short-run elasticity for gasoline has been estimated to be 0.2, while the long-run elasticity has been estimated at 0.7. In the short run, the only options consumers have in response to high gas prices are to take fewer trips and to use public transportation. But if gasoline prices stay high, consumers will start buying higher mile-per-gallon cars, depressing overall demand for gasoline even further. A retailer raising the price of milk by 20 cents may not see much change in milk sales for the first week or so and conclude that the price elasticity of milk is low. But he will likely see a much greater deterioration in demand over time. The reason is that customers who come to shop for milk after the price rise will still buy milk, since it is too much trouble to go to another store. But some customers will note the higher price and switch stores the next time they shop.

On the other hand, the long-run price elasticity of many durable goods—such as automobiles and washing machines—is lower than the short-run elasticity. The reason is that customers initially respond to a price rise by postponing the purchase of a new item. However, they will still purchase at some time in the future, so the long-run effect of the price change is less than the short-run effect.

Finally, it is important to specify the level at which we are calculating elasticity. Market elasticity measures total market response if all suppliers of a product increase their prices—perhaps in response to a common cost change. Market elasticity is generally much lower than the price-response elasticity faced by an individual supplier within the market. The reason is simple: If all suppliers raise their price, the only alternative faced by customers is to purchase a substitute product or to go without. On the other hand, if a single supplier raises its price, its customers have the option of defecting to the competition.

Table 3.1 shows some elasticities that have been estimated for various goods and services. Note that a staple such as salt is very inelastic—customers do not change the amount of salt they purchase very much in response to *market* price changes. On the other hand, we would expect that price elasticity of the market-response function faced by any individual seller of salt to be quite large—since salt is a fungible commodity in a highly competitive market.

TABLE 3.1
Estimated price elasticities for various goods and services

Good	Short-run elasticity	Long-run elasticity
Salt	0.1	—
Airline travel	0.1	2.4
Tires	0.9	1.2
Restaurant meals	2.3	—
Automobiles	1.2	0.2
Chevrolets	4.0	

This effect can be seen in Table 3.1 in the difference between the short-run elasticity for automobile purchases (1.2) and the much larger elasticity (4.0) faced by Chevrolet models. The table also illustrates the fact that long-run elasticity is greater than the short-run elasticity for airline travel (where customers respond to a price rise by changing travel plans in the future and traveling less by plane), but the reverse is true for automobiles (where consumers respond to price rises by postponing purchases).

3.1.3 Price Response and Willingness to Pay

So far, we have treated the price-response function as simply given. In reality, demand is the result of thousands, perhaps millions, of individual buying decisions on the part of potential customers. Each potential customer observes our price and decides whether or not to buy our product. Those who do not buy our product may purchase from the competition, or they may decide to do without. The price-response function specifies how many more of those potential customers would buy if we lowered our price and how many current buyers would not buy if we raised our price. Thus the price-response function is based on assumptions about customer behavior. We usually cannot directly track the thousands or even millions of individual decisions that ultimately manifest themselves in demand for our product.⁸ However, it is worthwhile to understand the assumptions about customer behavior that underlie the price-response functions so that we can judge if the price-response function is based on assumptions appropriate for the application. The most important of such models of customer behavior is based on *willingness to pay*.

The willingness-to-pay approach assumes that each potential customer has a *maximum willingness to pay* (sometimes called a *reservation price*) for a product or service. A customer will purchase if and only if the price is less than her maximum willingness to pay. (We will use *willingness to pay*, sometimes abbreviated *w.t.p.*, to mean “maximum willingness to pay.”) For example, a customer with a willingness to pay of \$253 for an airline ticket from New York to Miami will purchase the ticket if the price is less than or equal to \$253 but not if it is \$253.01 or more. In this case, $d(253)$ equals the number of customers whose maximum willingness to pay is at least \$253. A customer with a maximum willingness to pay of \$0 (or less) will not buy at any price.

Define the function $w(x)$ as the w.t.p. distribution across the population. Then, for any values $0 \leq p_1 < p_2$:

$$\int_{p_1}^{p_2} w(x) dx = \text{fraction of the population that has w.t.p. between } p_1 \text{ and } p_2$$

We note that $0 \leq w(x) \leq 1$ for all nonnegative values of x . Let $D = d(0)$, the maximum demand achievable. Then we can derive $d(p)$ from the w.t.p. distribution from

$$d(p) = D \int_p^{\infty} w(x) dx \quad (3.6)$$

We can take the derivative of the corresponding price-response function to obtain

$$d'(p) = -Dw(p)$$

which is nonpositive, as required by the downward-sloping demand curve property. Conversely, we can derive the willingness-to-pay distribution from the price-response function⁹ using

$$w(x) = -d'(x)/d(0)$$

Example 3.5

The total potential market for a spiral-bound notebook is $D = 20,000$, and willingness to pay is distributed uniformly between \$0 and \$10.00 as shown in Figure 3.3. This means that

$$w(x) = \begin{cases} 1/10 & \text{if } 0 \leq x \leq \$10 \\ 0 & \text{otherwise} \end{cases}$$

We can apply Equation 3.6 to derive the corresponding price-response curve:

$$\begin{aligned} d(p) &= 20,000 \int_p^{10} (1/10) dx \\ &= 20,000(1 - p/10) \\ &= 20,000 - 2,000p \end{aligned}$$

The price-response curve $d(p) = 20,000 - 2,000p$ is a straight line with $d(0) = 20,000$ and a satiating price of \$10.00.

Example 3.5 illustrates a general principle:

A uniform willingness-to-pay distribution corresponds to a linear price-response function, and vice versa.

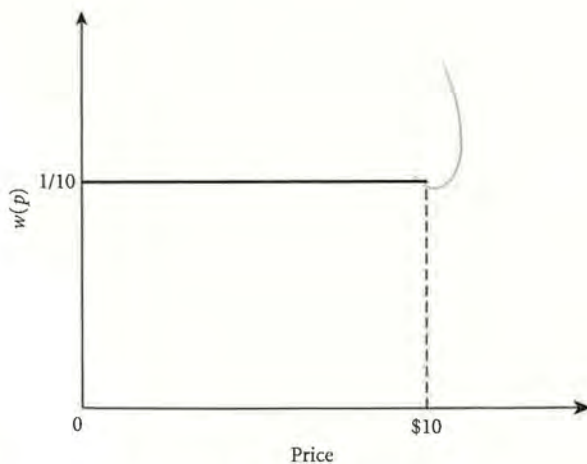


Figure 3.3 Uniform willingness-to-pay distribution.

One of the advantages of Equation 3.6 is that it partitions the price-response function into a total-demand component D and a willingness-to-pay component $w(x)$. This is often a convenient way to model a market. For example, we might anticipate that total demand varies seasonally for some product while the willingness-to-pay distribution remains constant over time. Then, given a forecast of total demand $D(t)$ for a future period t , our expected price-response function in each future period would be

$$d(p(t)) = D(t) \int_p^P w(x) dx \quad (3.7)$$

This approach allows us to decompose the problem of forecasting total demand from the problem of estimating price response. It also allows us to model influences on willingness to pay and total demand independently and then to combine them. For example, we might anticipate that a targeted advertising campaign will not increase the total population of potential customers, $D(t)$, but that it will shift the willingness-to-pay distribution. On the other hand, if we open a new retail outlet, we might anticipate that the total demand potential for the new outlet will be determined by the size of the population served, while the willingness to pay will have the same distribution as existing stores serving populations with similar demographics.

Of course, a customer's willingness to pay changes with changing circumstances and tastes. The maximum willingness to pay for a cold soft drink increases as the weather gets warmer—a fact that the Coca-Cola company considered exploiting with vending machines that changed prices with temperature (see Chapter 12 for a discussion of the “temperature-sensitive vending machine” idea). Willingness to pay to see a movie is higher for most people on Friday night than on Tuesday afternoon. A sudden windfall or a big raise may increase an individual's maximum willingness to pay for a new Mercedes Benz. To the extent that such changes are random and uncorrelated, they will not effect the overall w.t.p. distribution, since increasing willingness to pay on one person's part will tend to be balanced by another's decreasing willingness to pay. On the other hand, systematic changes across a population of customers will change the overall distribution and cause the price-response function to shift. Such systematic changes may be due to seasonal effects, changing fashion or fads, or an overall rise in purchasing power for a segment of the population. These systematic changes need to be understood and incorporated into estimating price response and future price response.

A disadvantage of the willingness-to-pay formulation is that it assumes that customers are considering only a single purchase. This is a reasonable assumption for relatively expensive and durable items. However, for many inexpensive or nondurable items, a reduction in price might cause some customers to buy multiple units. A significant price reduction on a washing machine will induce additional customers to buy a new washing machine, but it is unlikely to induce many customers to purchase two. However, a deep discount on socks may well induce customers to buy several pairs. This additional induced demand is not easily incorporated in a willingness-to-pay framework—willingness-to-pay models are most applicable to “big ticket” consumer items and industrial goods.

3.1.4 Common Price-Response Functions

Linear price-response function. We have seen that a uniform distribution of willingness to pay generates a linear price-response function. The general formula for the linear price-response function is

$$d(p) = D - mp \quad (3.8)$$

where $D > 0$ and $m > 0$. $D = d(0)$ is the *demand at zero price*. The general linear price-response function is shown in Figure 3.4. The *satiating price*—that is, the price at which demand drops to zero—is given by $P = D/m$. The slope of the linear price-response function is $-m$ for $p < 0 < P$ and 0 for $p \geq P$. The elasticity of the linear price-response function is $mp/(D - mp)$, which ranges from 0 at $p = 0$ and approaches infinity as p approaches P , dropping again to 0 for $p > P$.

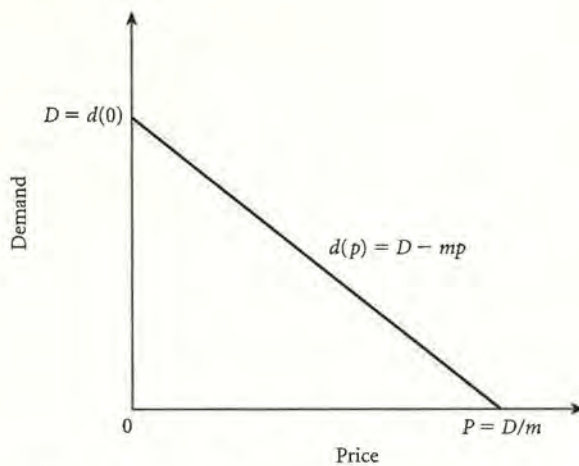


Figure 3.4 Linear price-response function.

We will use the linear price-response function in many examples because it is a convenient and easily tractable model of market response. However, it is not a realistic *global* representation of price response. The linear price-response function assumes that the change in demand from a 10-cent increase in price will be the same, no matter what the base price might be. This is unrealistic, especially when a competitor may be offering a close substitute. In this case, we would usually expect the effect of a price change to be greatest when the base price is close to the competitor's price.

Constant-elasticity price-response function. As the name implies, the constant-elasticity price-response function has a point elasticity that is the same at all prices. That is,

$$d'(p)p/d(p) = -\epsilon \quad \text{for all } p > 0 \quad (3.9)$$

where $\epsilon > 0$ is the elasticity. The price-response function corresponding to Equation 3.9 is

$$d(p) = Cp^{-\epsilon} \quad (3.10)$$

where $C > 0$ is a parameter chosen such that $d(1) = C$. For example, if we are measuring price in dollars and $d(\$1.00) = 10,000$, then $C = 10,000$. The slope of the constant-elasticity price-response function is

$$d'(p) = -C\epsilon p^{-(\epsilon+1)}$$

which is less than zero, confirming that constant-elasticity price-response functions are downward sloping. Some examples of constant-elasticity price-response functions are shown in Figure 3.5. Note that constant-elasticity price-response functions are neither finite nor satiating. Demand does not drop to zero at any price, no matter how high, and demand continues to approach infinity as price approaches zero. For these reasons, constant elasticity is usually not a good *global* assumption for price response.

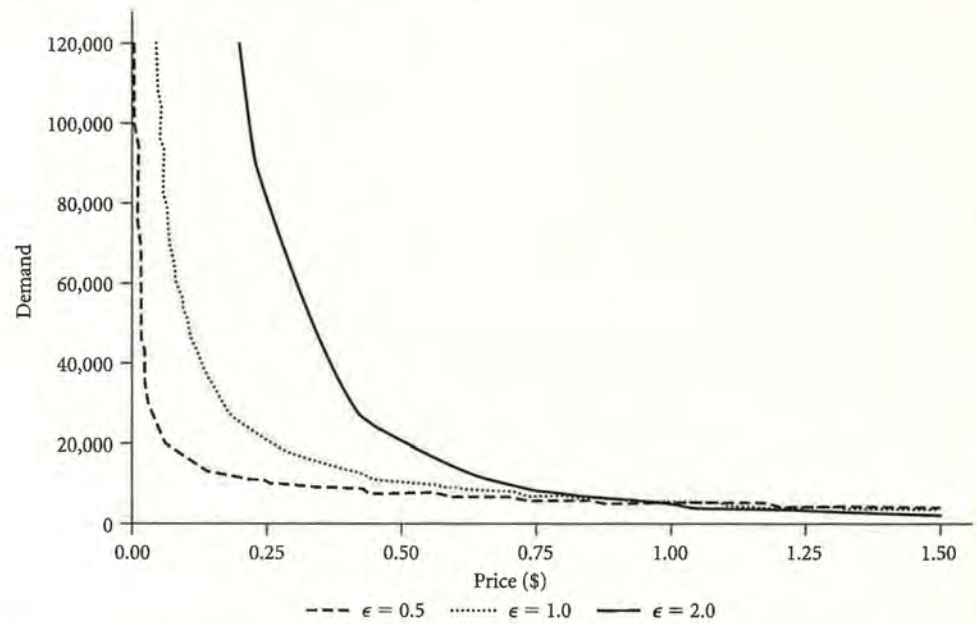


Figure 3.5 Constant-elasticity price-response functions.

It is interesting to analyze the impact of price on *revenue* for a seller facing constant-elasticity price response. Let $R(p)$ denote total revenue at price p . Then $R(p) = p d(p)$ and, for the constant-elasticity price-response function given by Equation 3.10,

$$\begin{aligned} R(p) &= pCp^{-\epsilon} \\ &= Cp^{(1-\epsilon)} \end{aligned}$$

Taking the slope of this function gives

$$\begin{aligned} R'(p) &= (1 - \epsilon)Cp^{-\epsilon} \\ &= (1 - \epsilon) d(p) \end{aligned}$$

Since $d(p) > 0$, the direction of the slope is determined by the $(1 - \epsilon)$ term. If $\epsilon < 1$ (that is, if demand is inelastic), then $R'(p) > 0$. This means that the seller facing constant *inelastic* price response can always increase his revenue by increasing price. In fact, he can also increase total operating margin by increasing price, so we would never expect to find a company maximizing operating margin setting its price such that the local elasticity is less than 1. If $\epsilon > 1$, a seller can increase revenue by decreasing price. In fact, the seller can maximize revenue by setting the price as close to zero—without actually becoming zero—as possible! If $\epsilon = 1$, then $R(p) = C$ for any value of p —revenue will not change at all as price changes. Revenue as a function of price for three different constant-elasticity price-response cases is illustrated in Figure 3.6.

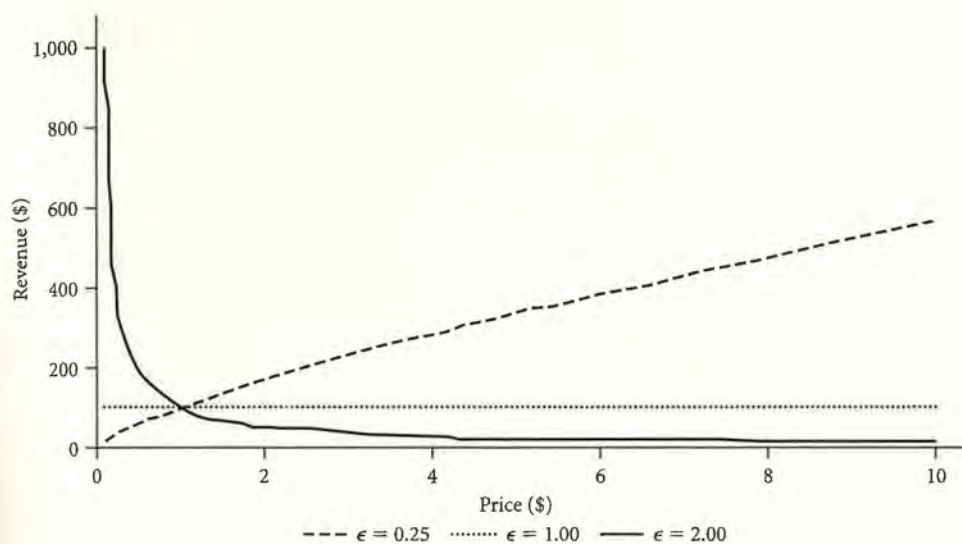


Figure 3.6 Revenue as a function of price for the constant-elasticity price-response function.

These properties make the constant-elasticity function problematic as a global price-response function. It is certainly the rare market in which a seller can maximize his revenue by either increasing his price toward infinity or dropping it as close to zero as possible. For these reasons, constant elasticity is a better representation of local price response than for use as a global price-response function. In real-world situations, we would expect price elasticity to change as price changes.

The constant-elasticity price-response function in Equation 3.10 corresponds to the willingness-to-pay distribution given by

$$w(x) = \epsilon x^{-(\epsilon+1)}$$

This distribution is given in Figure 3.7 for different values of ϵ . The figure shows that the constant-elasticity price-response function corresponds to a distribution of willingness to pay that is highly concentrated near zero.

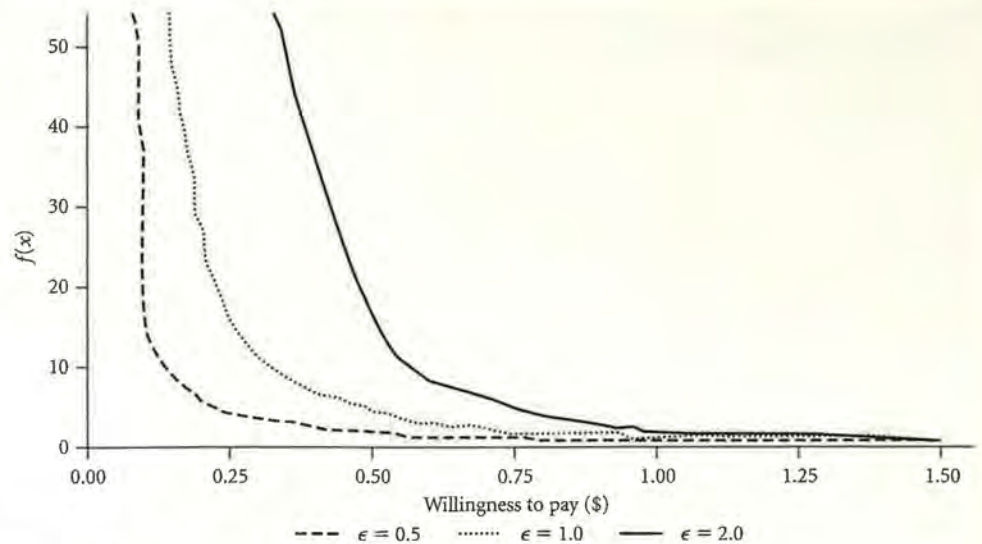


Figure 3.7 Constant-elasticity willingness-to-pay distributions.

The logit price-response functions. Both the linear and the constant-elasticity price-response functions are useful local models of price response. However, they have rather severe limitations as global models of customer behavior. This can be seen by considering their corresponding willingness-to-pay distributions. The linear price-response function assumes that w.t.p. is uniformly distributed between 0 and some maximum value. The constant-elasticity price-response function assumes that the distribution of willingness to pay drops steadily as price increases, approaching, but never reaching, zero. Neither of these functions would seem to represent customer behavior very realistically.

How would we expect customers to react to changes in price? Consider the case where we are selling a compact car. Assume that competing models are generally similar to ours and are selling at a market price of about \$13,000. If we price well above the market price, say, at \$20,000, we won't sell many cars. But the few customers who are buying from us at \$20,000 must be very loyal (or very ignorant)—we would not expect to lose very many more of them if we increased our price from \$20,000 to \$21,000. In other words, the elasticity of the price-response function is low at this price. On the other hand, if we priced well below the market price—say, at \$9,000—we would expect to sell lots of cars. At this price, almost anyone who wants to buy a compact car and is not incredibly loyal to another brand (or incredibly averse to ours) would be buying from us. Lowering our price another \$1,000 is unlikely to attract many new customers. Again, price elasticity is low.

What happens when we are close to the market price? In this case, we would expect small changes in our price to lead to substantial shifts in demand. There are a lot of customers who are more or less indifferent between our offering and that of the competition. Many of these customers can be persuaded to purchase our car if we ask only \$250 less than the market price, and even more will shift if we ask \$500 less than the market price. Similarly, if we charge \$250 above the market price we are likely to lose a lot of demand to the competition. In other words, price elasticity is highest when we are at (or near) the market price.

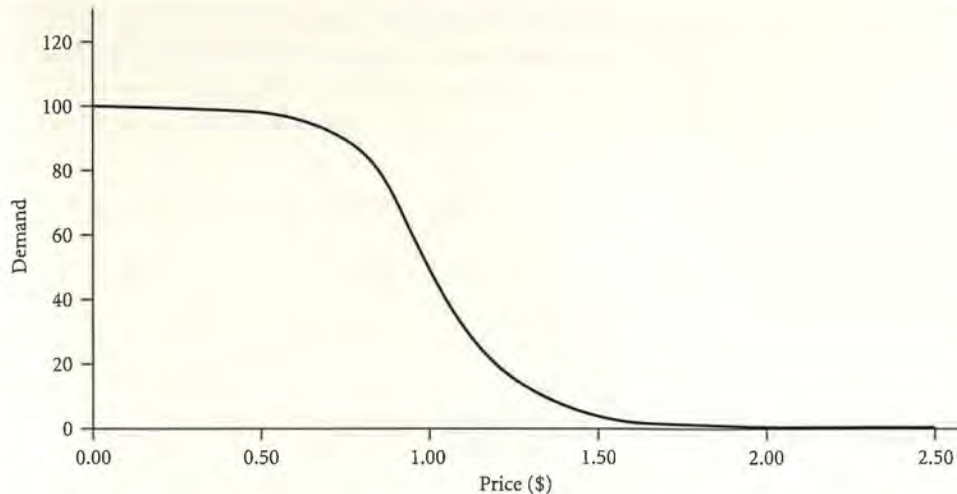


Figure 3.8 (Reverse) S-shaped, or sigmoid, price-response curve.

This kind of consumer behavior generates a response curve of the general form shown in Figure 3.8. When we price very low, we receive lots of demand, but demand changes slowly as we change price. In the area of the “market price,” demand is very sensitive to our price—small changes in price can lead to substantial changes in demand. At high prices, demand is low and changes slowly as we raise prices further. The price-response curve has a sort of reverse S shape. Empirical research has shown that this general form of price-response curve fits observed price response in a wide range of markets.

The most popular price-response function of this form is the *logit* price-response function:

$$d(p) = \frac{Ce^{-(a+bp)}}{1 + e^{-(a+bp)}} \quad (3.11)$$

Here, a , b , and C are parameters with $C > 0$ and $b > 0$. a can be either greater than or less than 0, but in most applications we will have $a > 0$. Broadly speaking, C indicates the size of the overall market and b specifies price sensitivity. Larger values of b correspond to greater price sensitivity. The price-sensitivity curve is steepest at the point $\hat{p} = -(a/b)$, as indicated in Figure 3.8. This point can be considered to be approximately the “market price.”

The logit price-response curve is shown for different values of b in Figure 3.9. Here we have fixed $\hat{p} = \$13,000$, so $a = -\$13,000 \times b$ for each curve. Demand is very sensitive to price when price is close to $\hat{p}$. Higher values of b represent more price-sensitive markets. As b grows larger, the market approaches perfect competition. In other words, the price-response curve increasingly approaches the perfectly competitive price-response function in Figure 3.1.

Some of the characteristics of the logit price-response function are shown in Table 3.2. Logit willingness to pay follows a bell-shaped curve known as the *logistic distribution*. The logistic distribution is similar to the normal distribution, except it has somewhat “fatter tails”—that is, it approaches zero more slowly at very high and very low values. An example of the logistic w.t.p. distribution is shown in Figure 3.10. The highest point (mode) of the logistic willingness-to-pay distribution occurs at $\hat{p} = -(a/b)$, which is also the point at which

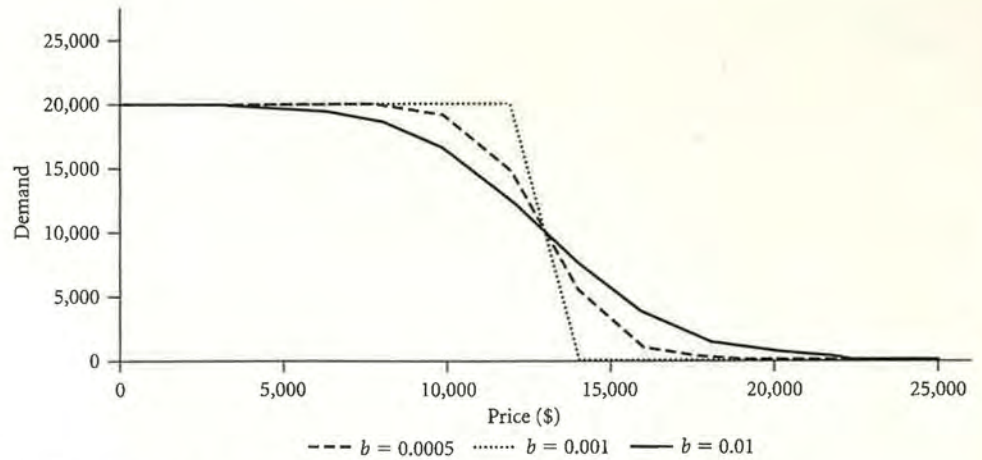


Figure 3.9 Logit price-response functions with $C = 20,000$ and $p = \$13,000$.

TABLE 3.2
Properties of the logit price-response function

Demand at 0	$d(0) = \frac{Ce^{-a}}{1 + e^{-a}}$
Slope	$d'(p) = \frac{-Cbe^{-(a+bp)}}{(1 + e^{-(a+bp)})^2}$
Elasticity	$\epsilon(p) = \frac{bp}{1 + e^{-(a+bp)}}$
Willingness-to-pay distribution	$w(x) = \frac{Ke^{-(a+bx)}}{(1 + e^{-(a+bx)})^2}$

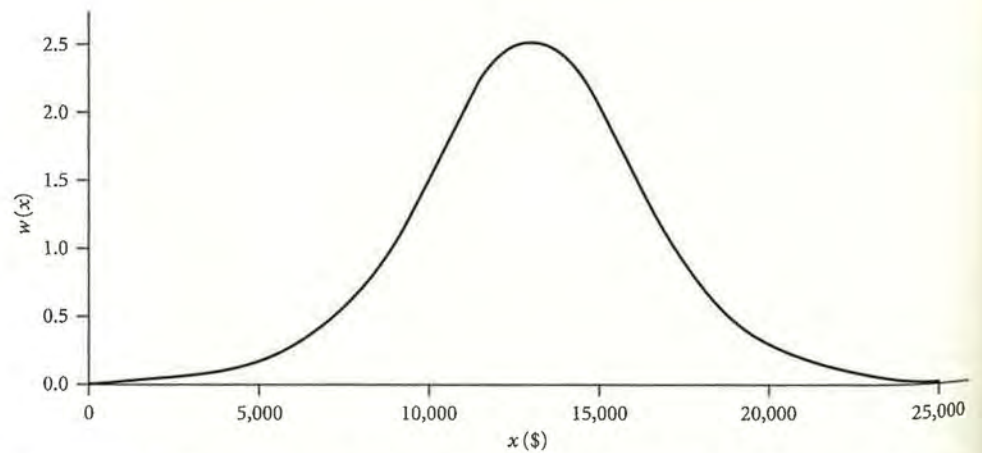


Figure 3.10 Willingness-to-pay distribution corresponding to the logit price-response function in Figure 3.9 with $b = 0.0005$. The y -axis has been scaled by a factor of 20,000.

the slope of the price response is steepest. This is a far more realistic w.t.p. distribution than those associated with either the constant-elasticity or linear price-response functions. For that reason, the logit is usually preferred to linear or constant-elasticity price-response functions when the effects of large price changes are being considered.¹⁰

3.2 PRICE RESPONSE WITH COMPETITION

So far we have not given much attention to the issue of competition. We have mentioned a “market price” as the steepest point in the logit price-response function, but we have not made an explicit link with competition. Yet competition is an important—some may say the most important—fact of life in any market of interest. Most managers identify competition as the number-one factor influencing their pricing. Many complain about being at the mercy of their competitors who have “stupid” or “irrational” pricing. How should we factor competition into our calculation of price response?

There are three different levels at which competition might be included in PRO. In increasing levels of sophistication and difficulty they are: incorporating competition in the price-response function, explicitly modeling consumer choice, and trying to anticipate competitive reaction.

3.2.1 Incorporating Competition in the Price-Response Function

In many cases, competitive prices may not be available at the time prices need to be optimized. This is especially true in business-to-business markets, when the pocket price of the customer may include discounts that are never made visible to the competition. But it is also true in many consumer markets. Retailers generally have a good idea of what their major competitors are charging at any particular time, but they usually don't have the time or resources to do exhaustive research on a daily or weekly basis.

Is there any hope of optimizing prices if we don't know what the competition is doing? The answer is yes. First of all, the price-response function we will be using in a market will be based on history. To this extent, *it already includes “typical” competitive pricing*. (Indeed, it would be difficult, if not impossible, to estimate a price-response function that did not include competition.) To the extent that our competitors will behave in a similar fashion in the future as they have in the past, the price-response function will be a fair representation of market response—including competitive response.

What if the market has changed? Then, in part, we must rely on the “fast-feedback loop” in the PRO process, as illustrated in Figure 2.5. If we have a competitor who responds more aggressively to a price action than in the past, this will be manifested in the demand we see in the next period. We have the opportunity to adjust our price-response function and re-optimize accordingly.

3.2.2 Consumer-Choice Modeling

Of course, if we have access to current competitive prices, we should be able to use that information to help us better predict customer response to our own price. This information is increasingly available, particularly for online markets. For example, the Online Petroleum

Information System (OPIS) posts a daily list of prices for different grades of petroleum products offered by different sellers in all major markets in the United States and Canada. How can this type of information be used by sellers in setting their prices?

Consumer-choice modeling is based on the situation in which a number of competitors are providing similar products or services to some population of customers. Each of the competitors sets a price for the product. Each customer has a vector of willingnesses to pay (or reservation prices), one for each product. These different w.t.p.'s represent the different values that she places on the features and "brand value" associated with each product—a more desirable product will have a higher associated w.t.p. The *surplus* for each customer and each product is defined as the difference between the customer's willingness to pay and the price of each product. Each customer purchases the product with the highest positive surplus. If none of the products has a positive surplus for a customer, she will not purchase at all.

Example 3.6

A customer is shopping for stereo amplifiers at the local Stereo Hut store. Stereo Hut sells five amplifiers at the following prices:

Model	Price	Willingness to pay	Surplus
Audio Two	\$400	\$250	-\$150
Koshiba	\$650	\$950	\$300
Takata and Fuji	\$500	\$550	\$50
Soundmaster	\$750	\$950	\$200
Cacophonia	\$1,200	\$1,300	\$100

This customer would purchase the Koshiba amplifier since it provides her with the highest surplus.

Note that, in the example, the customer does not buy either the cheapest alternative (Audio Two) or his favorite alternative (Cacophonia). Rather, she purchases the alternative that yields the highest surplus. We would expect that each customer would have a different set of reservation prices based on the value she places on the various features of each alternative or her preference for a particular brand. As before, we can imagine that the reservation prices would be distributed across the population in an n -dimensional bell-shaped curve. Assume now we have a fixed population of potential buyers and that each of these buyers has positive surplus for at least one of the alternatives. Then, the market share that will be obtained by a particular alternative is

$$\text{Market share of alternative } i = \text{Fraction of buyers for whom} \\ w_i - p_i > w_j - p_j \text{ for all } j$$

The idea behind consumer-choice modeling is to predict the share of each alternative given all of the prices. Let $p = (p_1, p_2, \dots, p_n)$ be the vector of prices for the alternatives

and let μ_i be the market share of alternative i . Then, a market-share function determines $\mu_i = f_i(p)$ for all i . A market-share function will have the following characteristics:

- $0 \leq f_i(p) \leq 1$. The market share of each alternative is between 0 and 1.
- $\sum_{i=1}^n f_i(p) = 1$. Every buyer chooses some alternative.
- $\partial f_i(p) / \partial p_i < 0$. Increasing the price of a product decreases its market share.
- $\partial f_i(p) / \partial p_j > 0$ for $i \neq j$. Increasing the price of a product increases the market share of other products. (The products are substitutes.)

Given a market-share function and a total demand D , the demand for product i is given by

$$d_i(p) = D\mu_i = Df_i(p) \quad (3.12)$$

The multinomial logit. The most widely used consumer choice model is the *multinomial logit* (MNL), defined by

$$\mu_i(p) = \frac{e^{-b_i p_i}}{\sum_{j=1}^n e^{-b_j p_j}} \quad (3.13)$$

where $b_j > 0$ is a parameter that varies by alternative. It is easy to see that $\mu_i(p)$ in Equation 3.13 satisfies all the conditions for a valid market-share function. The b_j are a measure of price responsiveness for each alternative—a high value of b_j means that alternative j is highly price sensitive, while a lower value of b_j indicates a lower value of price sensitivity.

Example 3.7

Statistical analysis of historical purchases has determined that the MNL provides a good fit to customer choices at Stereo Hut. The values of b_j for each amplifier are shown here along with the predicted market share at the current prices.

Model	Price	b_j	Market share
Audio Two	\$400	2.5	0.31
Koshiba	\$650	3.077	0.11
Takata and Fuji	\$500	2.4	0.25
Soundmaster	\$750	2.4	0.14
Cacophonia	\$1,200	1.25	0.19

Note that the Koshiba amplifier has a high level of price sensitivity and achieves a market share considerably lower than the Cacophonia, even though its price is almost 50% lower.

Relation of the multinomial logit to the logit price-response function. You may have guessed that there must be some relationship between the MNL and the logit price-response function. And indeed there is: *In the case where competitive prices are constant, the MNL reduces to the logit price-response function.* To see this, assume that prices $p_2, p_3, \dots, p_n$ are

held constant. Then we can replace the competitive-price term in Equation 3.13 with a constant. That is, we can set

$$K = \sum_{i=2}^n e^{-b_i p_i}$$

giving us

$$d_1(p_1) = \frac{De^{-b_1 p_1}}{e^{-b_1 p_1} + K} = \frac{De^{-b_1 p_1}/K}{e^{-b_1 p_1}/K + 1} = \frac{De^{-b_1 p_1} e^{-(\ln K)}}{e^{-b_1 p_1} e^{-(\ln K)} + 1} = \frac{De^{-(a+b_1 p_1)}}{e^{-(a+b_1 p_1)} + 1} \quad (3.14)$$

where we have set $a = \ln(K)$ and we use the fact that $\ln(1/K) = -\ln(K)$. This is the same as the logit price-response function specified in Equation 3.11. If we anticipate that competitive prices will be largely stable, the MNL provides little or no additional predictive value over a logit price-response function.

Strengths and weaknesses of consumer-choice modeling. There is a vast literature on consumer-choice modeling, and the issues involved with estimating the parameters of consumer-choice models have been much studied. Statistical software packages such as SAS include procedures for estimating the parameters of the logit or probit market-share functions. This is a great strength of the consumer-choice approach to price-response estimation.

There are a number of weaknesses to consumer-choice modeling as well. One is that, at least as we have posed it, the models assume that all customers purchase some alternative. However, it may be that some customers choose not to purchase at all, since their willingness to pay for every alternative may be below the price of that alternative. Thus, the estimation of the total market D is not really independent of the prices being offered. Indeed, it seems intuitive that an aggressive discount by one supplier would not only siphon customers away from competitors but actually induce customers into the market who might not have purchased from any alternative.

Another drawback of the consumer-choice modeling approach is that, in theory, it requires information on all competitive prices. However, in most markets, companies typically identify two to five companies they consider important competitors. For example, Table 11.2 lists the annual percentage rates (APRs) for unsecured consumer lenders for 16 different banks in the United Kingdom. While all of these 16 banks (and even more) compete in some sense for the loan business, in reality any lender would only consider four or five of them to be important competition. Other surveys have shown that in many consumer markets three or four major brands dominate a category, although there may be scores of smaller competitors. In these cases, one commonly used approach is to derive a *competitive index price* by weighting the prices offered by the major competitors and using this index as a single competitive price in a multinomial logit.

3.2.3 Anticipating Competitive Response

Even extensive consumer-choice modeling would seem to take us only partway to Nirvana. After all, if we are taking competitive prices into account in setting our prices, we should anticipate that our competitors will take our price into account when they set their prices. If we drop a price, we should anticipate the possibility that competitors will match, possibly eras-

ing much of the additional demand we might otherwise predict. The results of raising a price would certainly be different depending upon whether or not our competitors decided to match.

Attempts to predict competitive response and incorporate it into current pricing decisions falls within the realm of decision analysis, or game theory. While these approaches have their applications to strategic pricing, they are far less relevant to the tactical decisions of pricing and revenue optimization. For example, there is a vast literature on the use of game theory in pricing. This literature has little application to the day-to-day tactical pricing issues that fall within the scope of pricing and revenue optimization. There does not appear to be a single pricing and revenue optimization system that explicitly attempts to forecast competitive response using game theory as part of its ongoing operation. There are many reasons for this, but I believe there is one that is particularly important. Pricing and revenue optimization is based on playing the “best response” (i.e., finding the expected contribution-maximizing price) to whatever the competitors are currently doing. Theoretical examinations of markets find that this is almost always a good thing to do and often the best possible that can be done. For example, if customers in a market choose a supplier based on the multinomial logit price-response model described in Section 3.2.2, then the best each supplier can do is to set the price that maximizes expected contribution given the prices set by the competition. More sophisticated strategies cannot yield more.¹¹

The philosophy of pricing and revenue optimization is to make money by many small adjustments, searching for and vacuuming up small and transient puddles of profit as they appear in the marketplace. It is about making a little more money from each transaction. It is not about finding the great pricing move that will stagger the competition with one blow. Many of the price adjustments called for by pricing and revenue optimization are likely to fall below the radar screen of the competition and may not trigger any explicit response whatsoever.

This does not mean that competitive response need never be considered in pricing. The potential competitive responses to major “bet-the-company” bids or substantial changes in pricing strategy need to be carefully considered. For example, in the early 1990s, Hertz Rent-a-Car initiated the use of a sophisticated pricing optimization system. Previously, Hertz, like other national rent-a-car companies, had changed prices only rarely. Whenever one of the rental car companies dropped a price, the others were likely to follow immediately—often with even larger drops—precipitating an industrywide fare war. Hertz took great pains to communicate to the industry that its new system would be changing prices much more commonly than before—some prices would go up, some would go down, but all would change much more frequently. The communication was successful: Hertz was able to initiate its revenue management program without inciting retaliatory price wars. Once Hertz initiated revenue management, it was able to generate additional revenue through thousands of small adjustments to prices that its competitors were unable to match.

3.3 INCREMENTAL COSTS

Costing seems like it should be the easy part of pricing and revenue optimization. After all, a company may not know the price-response function it faces, but surely it should know its

own costs. Unfortunately, things are not quite so simple. With any exposure to the field of accounting one will quickly realize that the simple question "How much did this product cost?" usually does not have a simple answer. Rather, the "cost" of producing even a simple product depends on who wants to know and what they are going to use the answer for. Accountants have derived a host of product-costing methodologies, including fully allocated costs, partially allocated costs, marginal costs, avoided costs, and financial costs, to name just a few.

While each costing methodology has its place, pricing and revenue optimization decisions are based on the *incremental cost* of a customer commitment, where incremental cost is defined as follows.

The incremental cost of a customer commitment is the difference between the total costs a company would experience if it makes the commitment and the total cost it would experience if it doesn't.

Calculating incremental cost for a customer commitment depends critically on the context and nature of the commitment. Here are some examples.

- An airline is considering whether or not to sell a single seat to a passenger on a future flight. The incremental cost is the additional meal and fuel cost that would be incurred from flying an additional passenger plus any commissions or fees the airline would pay for the booking.
- A retailer buys a stock of fashion goods at the beginning of the season. Once he has bought the goods, he cannot return them and cannot reorder. During the season he wants to set and update the prices that will maximize the total revenue he will receive from the fashion goods. Since the cost of the goods is sunk and selling a unit will not drive any additional future sales, his incremental cost per sale is zero.¹²
- A drugstore sells bottles of shampoo and replenishes its stock weekly. Because selling a unit will result in an additional unit order, the incremental cost for selling a bottle of shampoo is the wholesale unit cost.
- A distributor is bidding for a year-long contract with a hospital to be the preferred provider for disposable medical supplies. From previous experience, the distributor knows that this hospital is an expensive customer requiring high levels of customer support, wide variances in orders, and high rates of product returns. The expected incremental cost of the contract includes not only the expected cost of the items the hospital will purchase, but also the expected cost of customer service, operating costs and holding costs driven by the wide variance in orders, and the expected costs of returns.

These examples illustrate some of the key characteristics of incremental cost. These characteristics can be summarized as follows.

- *Incremental cost is forward looking.* It is based on the effect a customer commitment will have on future costs. Costs that have already been taken or that are driven by

decisions already made prior to the customer commitment are sunk and do not enter into the calculation of incremental cost.

- *Incremental cost is marginal.* It is the expected cost of making *this* customer commitment. The incremental cost of making this commitment may not be the same as the average cost of similar commitments made in the past.
- *Incremental cost is not fully allocated cost.* Only costs that change as the result of a customer commitment are part of the incremental cost. Overhead, or fixed costs of staying in business, may not be allocated to any specific commitment. As a result, the incremental cost of a customer commitment is usually less than the fully allocated cost. Further, the sum of the incremental costs of all commitments will be less than the total operating cost of the company, since there will remain an unallocated residual fixed cost after all the incremental costs have been totaled.
- *The elements of incremental cost can depend on the type, size, and duration of the commitment.* Specifically, if the customer commitment is a multiunit order, then order-based or setup costs need to be allocated across all the units in the order.
- *The incremental cost of a commitment can be uncertain.* Trucking companies such as Roadway Express and Yellow Freight sell contracts to shippers. Each contract covers the next year and commits the trucking company to carry all the freight tendered by the customer at an agreed-on tariff. The incremental cost associated with one of these contracts is likely to be highly uncertain at the time the commitment is made. First of all, the amount, timing, and origin and destination of the freight the customer will tender over the next year is uncertain. In addition, the cost of moving the customer's freight will depend on the freight the company will move for other customers, which is also uncertain. This cost will depend in part on what portion of the customer's freight is *backhaul* business, which can utilize excess capacity on existing trucks, and what portion will be *headhaul* freight, which will require running new trucks.

In each case, the calculation of incremental cost requires understanding the nature of the customer commitment and then estimating the additional costs that would be generated by making the commitment—or, equivalently, the costs that would be avoided by not making the commitment. The methodology behind calculating incremental costs is closely related to activity-based costing (ABC). Activity-based costing is a management accounting approach to allocating costs to their underlying causes in order to give a clearer view of the real sources of cost within an organization. A good introduction to activity-based costing is *Cost & Effect* by Kaplan and Cooper (1997).

3.4 THE BASIC PRICE OPTIMIZATION PROBLEM

The difference between the price at which a product is sold and its incremental cost is called its *unit margin* or just *margin*. The sum of the margins of all products sold during a time pe-

In this chapter we treat one of the most fundamental concepts in PRO—*price differentiation*. Price differentiation refers to the practice of a seller charging different prices to different customers, either for exactly the same good or for slightly different versions of the same good. Price differentiation is a powerful way for sellers to improve profitability. It also adds a new level of complexity to pricing, often creating a need to use analytical techniques to improve the calculation and updating of prices over time.

We use the term *price differentiation* to refer to the ways that additional profit can be extracted from a marketplace by charging different prices. Tactics for price differentiation include charging different prices to different customers (or groups of customers) for *exactly* the same product, charging different prices for different versions of the same product, and combinations of the two. The term *price discrimination* is used in the economics literature to refer to much the same thing. We use the term *price differentiation*, rather than the more common *price discrimination*, in part to avoid the negative connotations associated with the word *discrimination*. However, we also want to stress that price differentiation includes not only charging different prices to different customers for the same product (group pricing) but also the less controversial strategies of product versioning, regional pricing, and channel pricing.

There is both art and science to price differentiation. The art lies in finding a way to divide the market into different segments such that higher prices can be charged to the high-willingness-to-pay segments and lower prices to the low-willingness-to-pay segments. There is no one way to segment customers that applies to all possible markets. Instead, there is a variety of techniques that can be applied in different ways, depending on the characteristics of a market, the competitive environment, and the character of the goods or services being sold. The science lies in setting and updating the prices in order to maximize overall return from all segments.

We start by using a simple example (based on the widget maker from Chapter 3) to illustrate the economic principles behind price differentiation.

4.1 THE ECONOMICS OF PRICE DIFFERENTIATION

Let us return to the case of the widget maker in Section 3.4.2. The widget maker, looking for a way to increase his profitability, might start by taking a close look at his price-response function, as illustrated in Figure 3.13. He would see that 10,000 customers are willing to pay some amount greater than zero for his widget, 5,000 are willing to pay \$6.25 or more, 3,000 are willing to pay \$8.75 or more (the contribution-maximizing price), and 1,200 are even willing to pay \$11.00 or more. The marketing department might see an opportunity to improve profitability. After all, there are 2,000 customers willing to pay \$10.00 or more who are purchasing at \$8.75. This is a good deal for these customers, but it is also at least \$1.25 in additional revenue “left on the table” from each sale. Furthermore, 3,000 customers are willing to pay more than the production cost (\$5.00) but less than \$8.75—the sales price. Each of these is a potentially profitable sale lost because the price is too high. What if the seller could determine the maximum amount that each customer would be willing to pay and could charge that amount to everyone willing to pay more than the per-unit cost of \$5.00? This would be pricing nirvana—the ultimate in “one-to-one” pricing.

The total potential opportunity for profit improvement from price differentiation is shown in Figure 4.1. With a single price, the widget maker will charge \$8.75 per unit and realize a total profit of $(\$8.75 - \$5.00) \times 3,000 = \$11,250$. This is the area of region A in Figure 4.1. “Perfect” price differentiation would enable the seller to charge each customer willing to pay more than \$8.75 per unit her exact willingness to pay. This would provide the additional revenue in region B in the figure. There is an additional opportunity to sell widgets to customers who are willing to pay more than \$5.00, the unit production cost, but less than \$8.75. This potential profit is shown as region C in the figure. The sum of the three regions is the total contribution that the widget maker would realize if he were able to charge every potential customer exactly at her willingness to pay. This is an upper bound on what

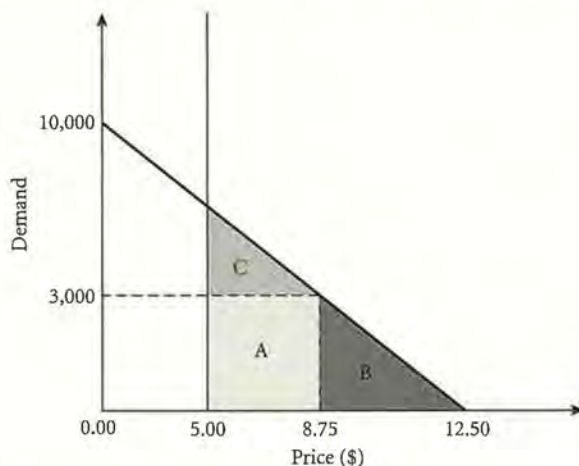


Figure 4.1 Contribution opportunity from price differentiation.

could possibly be realized under any price differentiation program. Charging every customer exactly her willingness to pay is known as *third-degree price discrimination* in the economics literature.

While it is unrealistic to assume that this total potential could ever be captured, the sheer magnitude of the potential gain means there is a powerful motivation for sellers to tailor different prices to different buyers according to their willingness to pay. Assume that the widget maker finds he can divide his market into two segments. One segment consists of all customers willing to pay more than \$7.00 for widgets. The other segment consists of all the customers willing to pay \$7.00 or less. The corresponding price-response curves are:

$$d_1(p_1) = \min[4,400, (10,000 - 800p_1)^+]$$

$$d_2(p_2) = (5,600 - 800p_2)^+$$

where $d_1(p_1)$ is the price-response curve for customers with a high willingness to pay and $d_2(p_2)$ is the price-response curve for customers with the lower willingness to pay. These two price-response curves are shown in Figure 4.2. The sum of these two curves is the original price-response curve in Figure 3.13. These two segments taken together make up the same total market that the widget maker faced before. The difference is that now he can offer a different price to each of the two segments. We assume for this example that the widget maker can perfectly identify customers as belonging to one group or the other and can then offer each customer the appropriate price, without any opportunity for resale or arbitrage between the two groups.

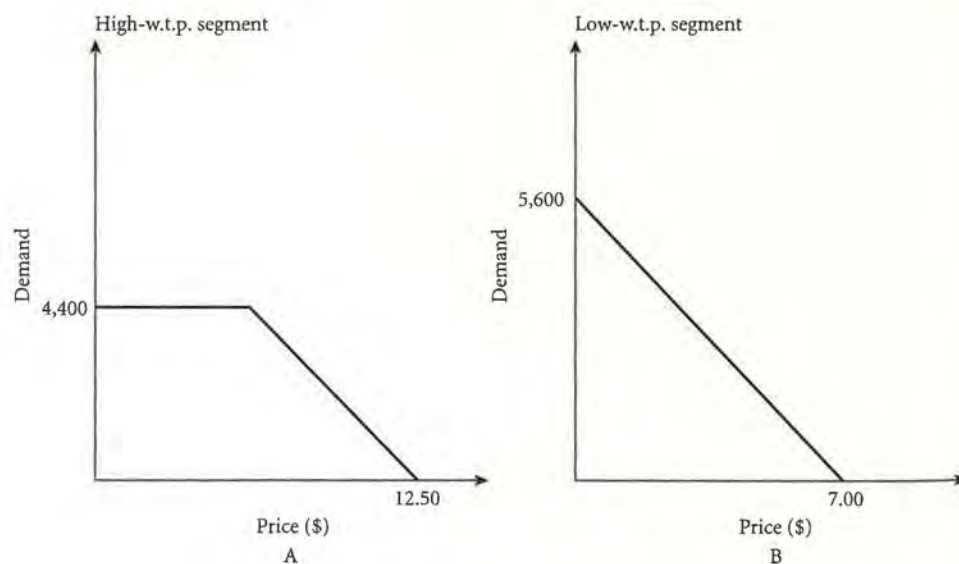


Figure 4.2 Price-response curves for two segments.

The widget maker can determine the optimal prices for each segment by solving Equation 3.18 twice—once for each segment. The results (along with a comparison to the unsegmented case) are shown in Table 4.1. Dividing his customers between those with w.t.p.

TABLE 4.1
Impact of market segmentation

	Unsegmented	Segment 1	Segment 2	Total	Change (%)
Average price	\$8.75	\$8.75	\$6.00	\$8.17	-6.6
Demand	3,000	3,000	800	3,800	26.7
Revenue	\$26,250	\$26,250	\$4,800	\$31,050	18.3
Gross contribution	\$11,250	\$11,250	\$800	\$12,050	7.0
Net contribution	\$1,250	—	—	\$2,050	85

greater than \$7.00 and those with w.t.p. less than \$7.00 and charging the optimal prices to the different segments enables the widget maker to increase his revenues by 18% and his total contribution by 7%. It is interesting to note that customers benefit as well. The same 3,000 customers who would buy under a single price at \$8.75 still get to buy at \$8.75. In addition, 800 additional customers get to purchase at the new low price of \$6.00. These 800 customers are priced out of the market if the seller can only charge a single price. In this case, price differentiation is a win-win situation since the seller is certainly better off and all of the buyers are at least as well off as before. However, price differentiation is not always such a boon for consumers, as we discuss in Section 4.6.

4.2 LIMITS TO PRICE DIFFERENTIATION

If price differentiation is such a powerful way for sellers to increase contribution, why don't all of them do it? The reason is that there are powerful real-world limits to price differentiation.

1. *Imperfect segmentation.* The brain-scan technology required to determine the precise willingness to pay of each customer has not yet been developed. The best that can be done is to create market segments such that the *average* willingness to pay is different for each segment.
2. *Cannibalization.* Under differential pricing, there is a powerful motivation for customers in high-price segments to find a way to pay the lower price. In the widget example, there is a strong motivation for high-willingness-to-pay customers who are being charged \$8.75 per widget to "masquerade" as low-willingness-to-pay customers and pay only \$6.00 per widget.
3. *Arbitrage.* Price differentials create a strong incentive for third-party *arbitrageurs* to find a way to buy the product at the low price and resell to high-w.t.p. customers below the market price, keeping the difference for themselves.

The presence of any one of these factors can eliminate the benefits of price differentiation.

Example 4.1

Assume that the widget seller has segmented his market and is charging the two prices shown in Table 4.1. If 300—that is 10%—of the high-willingness-to-pay customers are able to find a way to purchase widgets at the lower cost of \$6.00, it would totally eliminate the benefits of price differentiation.

The deterioration of benefits in Example 4.1 could arise from imperfect segmentation, cannibalization, arbitrage, or a combination of any of them.

4.3 TACTICS FOR PRICE DIFFERENTIATION

The previous section established two properties of price differentiation.

1. Price differentiation allows sellers to increase profitability by charging different prices to customers with different willingness to pay.
2. Cannibalization, imperfect segmentation, or arbitrage can destroy—or even reverse—the benefits of price differentiation.

The first property means that sellers have a tremendous incentive to find creative ways to position and price their products differently to different market segments. The second means that price differentiation needs to be carefully planned and managed in order to be successful. In this section we describe some of the most common and effective approaches to price differentiation used in different markets.

4.3.1 Group Pricing

Group pricing is the tactic of offering different prices to different groups of customers for exactly the same product.¹ The idea is to offer a lower price to customers with a low willingness to pay and a higher price to those with a high willingness to pay. In practice, “pure” group pricing requires determining whether or not a prospective buyer belongs to a particular group and using that information to determine which price to charge. Examples of pure group pricing include:

- Student discounts
- Senior citizen discounts
- “Ladies’ Night” specials
- Family specials
- Discounts for favored customers
- Favorable terms offered by manufacturers or wholesalers to large retailers such as Wal-Mart
- Lower prices offered to government, educational, and nonprofit organizations by suppliers

Four criteria must hold for group pricing to be successful.

- *There must be an unambiguous indicator of group membership.* Examples of such indicators include a student ID card or a driver’s license that lists age. Furthermore, it must be difficult or impossible for members of one group to masquerade as members of another. Otherwise, cannibalization could easily reduce or destroy the benefits of price differentiation.

- *Group membership must strongly correlate with price sensitivity.* Senior citizen discounts are predicated on the belief that senior citizens, *as a whole*, are more price sensitive than the public in general.
- *The product or service should not be easily traded or exchanged among purchasers.* This is necessary to avoid *arbitrage*—in which customers with access to low prices resell to customers who are quoted higher prices.
- *The segmentation must be both culturally and legally accepted.* While group pricing on the basis of age is broadly accepted, differentiating prices on the basis of other characteristics, such as race and gender, are controversial or illegal. The Robinson-Patman Act prohibits many forms of group pricing that wholesalers might want to use to charge differential prices to retailers.

Taken together, these criteria are so stringent that pure group pricing is relatively rare in direct consumer sales. It is most common in services. Services are often sold directly by the supplier so that Disneyworld, for example, can check a customer's age before selling a child's ticket. Furthermore, many services, such as health care and haircuts, are intrinsically non-transferable, so arbitrage is not an issue. Airlines originally began checking passenger identification not for security reasons but to prevent arbitrage.

Pure group pricing is also common in business-to-business sales. In Chapter 11 we discuss how businesses can estimate price responsiveness and develop customized prices for different business segments. Customized pricing in business-to-business sales is often combined with some form of product versioning, particularly in industries where individual orders are very complex and/or configurable.

4.3.2 Channel Pricing

Channel pricing is the practice of selling the same product for different prices through different distribution channels. Here are some examples.

- Barnes and Noble sells books for different prices online than through its outlets.
- Special "Web-only" fares for airline tickets are available through the Internet but not through travel agencies.
- Many fashion and home furnishing merchants offer lower prices for merchandise through mail-order catalogs than for the same merchandise sold through retail outlets.

As with other price differentiation schemes, there can be more than one reason why a seller might charge different prices through different channels. One is cost—for many companies, selling through the Internet is cheaper than selling through traditional channels. Continental Airlines calculated that distribution costs for a \$325 ticket are about \$43 when the ticket is sold through a travel agent versus only \$18 when the same ticket is sold through Continental's Web site. This difference in itself would be sufficient reason for Continental to charge lower prices on its Web site.

However, it is also the case that customers arriving via different channels have different price sensitivities. For personal loans, it has been shown that customers inquiring through the Internet are more price sensitive than those contacting a call center, who are in turn more price sensitive than those who apply for a loan at a retail branch. Furthermore, Internet customers who access a consolidator Web site or use a shopping 'bot tend to be more price sensitive than those who go directly to a bank's Web site. This is not surprising given the characteristics of the channels—it is generally easier and more convenient to shop and compare prices during a single Internet session than by making many phone calls. Thus differential willingness to pay is also a motivation for channel pricing.

4.3.3 Regional Pricing

Regional pricing is an extremely common price differentiation technique. Here are some examples.

- In Latin America, McDonalds sells hamburgers for higher prices in wealthy neighborhoods than in poorer ones.²
- A roundtrip New York–Tokyo ticket purchased in Japan will usually cost more for many airlines than the same ticket purchased in the United States.
- A glass of beer costs more at an airport bar than at the corner bar.

In each case, the price difference is based on the supplier's desire to exploit differences in price sensitivity between locations. After all, travelers at an airport are essentially a captive market and have few alternatives.

4.3.4 Couponing and Self-Selection

We have seen that group pricing is often both difficult and unpopular—difficult because it requires the seller to categorize customers on the basis of price sensitivity before quoting them a price, and unpopular because it often seems “unfair” to consumers. It is often much more convenient to differentiate prices in ways that allow customers to self-select. In a self-selection approach, both the list price and a discounted price are available to all customers, but it takes additional time, effort, or flexibility to obtain the discounted price. The idea is that those willing to make the additional effort to get the discount are generally more price sensitive than those who are not. Here are some examples.

- Retailers commonly offer discount coupons through newspapers, direct mail, and magazines.
- Retailers often offer mail-in rebates for purchasers of a good.
- Movie theaters charge lower prices for a weekday matinee than for a Saturday night show.
- Brand-name retailers such as Ralph Lauren, the Gap, and Liz Claiborne operate outlet stores in somewhat out-of-the-way locations in which merchandise is available for a substantial discount.

In Chapter 3 we saw how a seller can calculate the optimal single-period price when he has the freedom to produce (or order) as much of a good as he wants and he knows with certainty the price-response curve he faces. In this case the seller can maximize expected profit by charging the price at which his marginal cost equals his marginal revenue or, equivalently, the price at which his contribution margin ratio equals 1 over the elasticity. Unfortunately, things are a bit more complicated in the real world: Sellers have the freedom to adjust prices over many periods, they are unsure how customers will respond to different prices, and they face constraints on their ability to satisfy demand. Sellers need to be able to set prices in a dynamic, uncertain, and constrained world. In this chapter we focus on the influence of constrained supply on optimal prices.

When the supply of a good is constrained for any reason, the approaches presented in Chapter 3 do not necessarily give the price that maximizes net contribution. Return for a moment to the example of the widget maker. With only a single period to consider and with perfect knowledge of his market, the optimal actions for him were to set his price at \$8.75 and produce 3,000 units. In this case, he sells all 3,000 units and makes a total contribution of \$11,250. This solution assumes that supply is totally flexible *and* demands are deterministic *and* he is only going to operate during a single period. What should he do if he can only manufacture 2,800 widgets? What price should he set? That is the sort of question we address in this chapter.

We start by discussing the nature of supply constraints and the situations in which they occur. We then discuss how a seller can determine the optimal price to charge when faced by a supply constraint, and we introduce the important concept of *opportunity cost*. We extend the calculation of optimal prices to the case when a supplier has a segmented market and faces supply constraints. This leads to the tactic of variable pricing, which is used when a supplier has multiple units of constrained capacity and can change prices in order to balance supply and demand.

5.1 THE NATURE OF SUPPLY CONSTRAINTS

Most sellers face supply constraints at least some of the time. For example, most retailers replenish their stock of inventory at fixed intervals—often weekly. In between replenishment times, they are limited to selling their current inventory—in other words, they have a fixed supply. However, in most cases, retailers stock enough inventory that a stock-out is unlikely. A drugstore will typically have enough toothpaste, shaving cream, and aspirin in stock that it will not sell out except in cases of an extraordinary run on a particular item—bottled water prior to a hurricane, for example. In this case, the seller may technically have a limited amount of inventory on hand, but he does not need to consider a supply constraint when setting price. There are, however, many cases in which a seller needs to consider the constrained nature of supply in order to calculate the optimal price.

1. *Service providers* almost always face capacity constraints. A hotel is constrained by the number of rooms, a gas pipeline by the capacity of its pipes, and a barbershop by the number of seats (and barbers) it has available.
2. *Manufacturers* face physical constraints on the amount they can produce during a particular period. For example, Ford Motor Company in North America can produce about 475,000 vehicles per month.¹ In any particular month, Ford can only sell 475,000 vehicles plus whatever inventory it had on hand at the beginning of the month.
3. *Retailers and wholesalers* often sell goods that are not replenishable. These include fashion goods that are typically ordered once or electronic goods that are near the end of their life cycle.
4. *Intrinsically scarce or unique items*, such as beachfront property, flawless blue diamonds, van Gogh paintings, and Stradivarius violins, command premium prices because of their scarcity. In these cases, marginal cost is not an important determinant of price either because it is meaningless (as in the case of van Gogh paintings) or extremely low relative to the scarcity rent that sellers can command (as in the case of diamonds).

The treatment of a supply constraint depends on its nature. A *hard* constraint is one that cannot be violated at any price. Typically hotels and gas pipelines face hard constraints. On the other hand, air freight carriers often have the option to lease space on other carriers to carry cargo in excess of their capacity. The freight carrier's physical capacity is thus a *soft* constraint. Whether a supply constraint is hard or soft will, in part, depend on timing. An airline that learns two months in advance that it will be facing very high demand for a particular flight may be able to assign a larger aircraft. However, it may not have any option to increase capacity if it only learns about the high demand a week before departure.

5.2 OPTIMAL PRICING WITH A SUPPLY CONSTRAINT

In Section 3.4.2, the optimal unconstrained price of widgets was \$8.75, with corresponding demand of 3,000 units and a total contribution of \$11,250. This is optimal only if the seller is able to manufacture all 3,000 units—or order them from a third party. But what if he cannot satisfy all the demand at the optimal unconstrained price? Perhaps he has a supply chain bottleneck that limits his production. Perhaps he is a reseller with a fixed monthly quota he cannot exceed. In any case, the widget maker faces a rationing problem—he needs to determine which customers will be served and which will not. He has three basic options:

1. Do nothing—Keep the price at \$8.75 and let customers buy on a first-come/first-served basis until supply is exhausted.
2. Allocate the limited supply to favored customers.
3. Raise the price until demand falls to meet supply.

Alternatives two and three are not mutually exclusive—the seller could use a combination of allocations and price rises to manage the shortage. And, if he has segmented his market effectively, he could raise his average selling price by allocating most or all of the limited supply to higher-paying customers—this is the basic idea behind revenue management, which we discuss in Chapter 6.

If the widget maker cannot segment his market (allocating favored customers is a form of segmentation), his only options are to sell on a first-come/first-served basis or to raise the price. Let's assume he decides to raise the price in order to maximize his short-run contribution margin. Then he needs to solve the constrained optimization problem:

$$\max d(p)(p - c)$$

subject to

$$d(p) \leq b$$

where b is the maximum supply available. Let p^* be the optimal unconstrained price. If $d(p^*) \leq b$, the supplier doesn't need to do any further calculations: p^* is also the optimal constrained price. If, on the other hand, he finds that $d(p^*) > b$, then he needs to charge a *higher* price in order to maximize contribution. This situation is illustrated in Figure 5.1. Here, the capacity constraint is a horizontal line—the supplier can only produce quantities below that line. At the optimal unconstrained price, p^* , demand $d(p^*)$ exceeds capacity.

When the optimal unconstrained price generates demand that exceeds capacity, the supplier needs to calculate the *runout price*, that is, the price at which demand would exactly equal the supply constraint. In Figure 5.1, the runout price is the price $\hat{p}$ at which the price-response curve intersects the capacity constraint. In other words, it is the price at which $d(\hat{p}) = b$.

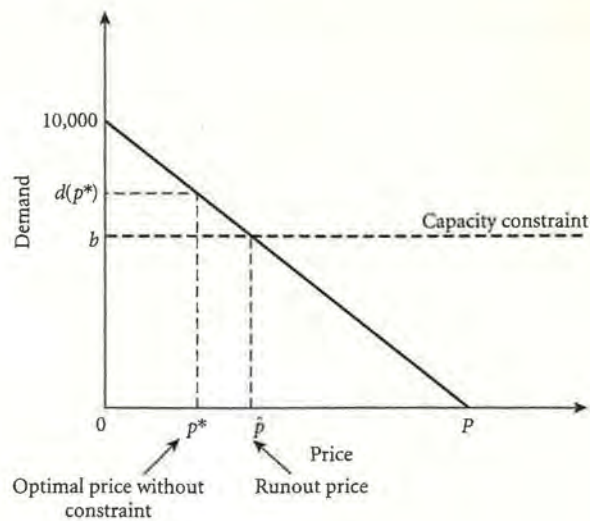


Figure 5.1 Pricing with a capacity constraint.

Example 5.1

Assume the widget seller faces the price-response curve $d(p) = 10,000 - 800p$ but can only supply a maximum 2,000 widgets during the upcoming week. Demand at the optimal unconstrained price of \$8.75 is 3,000 widgets. Therefore, the supply constraint is binding and he needs to price at the runout price. The runout price of $\hat{p} = \$10.00$ can be found by solving $d(\hat{p}) = 10,000 - 800\hat{p} = 2,000$.

The general principle behind calculating the optimal price with a supply constraint is the following.

The profit-maximizing price under a supply constraint is equal to the maximum of the runout price and the unconstrained profit-maximizing price. As a consequence, the profit-maximizing price under a supply constraint is always greater than or equal to the unconstrained profit-maximizing price.

The effects of different levels of capacity constraint on price and total revenue are shown in Table 5.1. For a binding capacity constraint, $b < 3,000$, the runout price can be found by inverting the price-response function:

$$\hat{p} = d^{-1}(b) = (10,000 - b)/800 \quad \text{for } b < 3,000$$

Table 5.1 shows the results of reducing capacity at increments from 3,000 units per month to 1,000 units per month. The optimal price increases steadily as the capacity decreases. The benefits of reoptimizing price in response to constrained capacity also increases steadily. As capacity becomes increasingly constrained, there is more and more pressure on the seller to increase the price.

5.5 VARIABLE PRICING

Consider the following pricing tactics.

- Most local telephone companies charge different rates for calls by time of day. For example, one plan might charge 3 cents per minute from 6:00 AM to 6:00 PM, 2 cents per minute from 6:00 PM to 11:00 PM, and 1 cent per minute from 11:00 PM to 6:00 AM.
- The San Francisco Opera charges a lower price for weeknight performances than for weekend performances. The most expensive box seats cost \$175 for a Wednesday performance and \$195 for a Saturday performance, while the least expensive balcony side tickets cost \$25 for a Wednesday performance and \$28 for a Saturday performance.
- The Colorado Rockies baseball team uses the four-tier pricing system for their home games shown in Table 5.2. Lower prices for games in April, May, and September are justified by the fact that the weather is more likely to be bad during these months, resulting in lower demand.³

TABLE 5.2
2002 Home ticket prices for the Colorado Rockies baseball team

Tier	Description	Average price
Value	Weekday games against low-drawing opponents during April, May, and September	\$16.97
Standard	Weekday games against division rivals during April, May, and September	\$18.00
Premium	Weekend games in May, June, July, and August	\$20.07
Classic	Opening day, games against the N.Y. Yankees, fireworks games	\$21.58

- The Gulf Power Company is an electric utility serving more than 370,000 retail customers in northwestern Florida. The majority of these retail customers pay a flat rate of \$0.057 per kilowatt hour for electricity. However, Gulf Power offers residential customers the alternative of purchasing their electricity under their Residential Select Variable Pricing (RSVP) program. Under the RSVP program, each weekday is divided into an off-peak, a shoulder, and an on-peak period. Electricity costs \$0.035 per kilowatt hour (kWh) during the off-peak period, \$0.046 during the shoulder period, and \$0.093 during the on-peak period. In the summer the on-peak period is from 11:30 AM to 8:30 PM, while in the winter the on-peak period is from 6:00 AM to noon. The summer off-peak period is from 10:30 PM to 6:00 AM and the winter off-peak period is from 10:00 PM to 5:30 AM. All other hours are shoulder period. In addition to these three periods, with a half hour's notice Gulf Power can declare a "critical period" during which electricity costs \$0.290 per kWh. Under the RSVP program, Gulf Power can declare up to 88 hours of critical period per year.⁴

What do these four pricing tactics have in common? They are all cases in which a seller with constrained capacity adjusts prices in response to anticipated changes in demand in

order to maximize return from fixed capacity. We call this tactic *variable pricing*. Industries where variable pricing is employed share four characteristics:

- Demand is variable but follows a predictable pattern.
- The capacity of a seller is fixed in the short run (or is expensive to change).
- Inventory is perishable or expensive to store—otherwise buyers would learn to predict the variation in prices and stockpile when the price is low.
- The seller has the ability to adjust prices in response to supply/demand imbalances.

Under these conditions, sellers can use variable prices to shape demand to meet fixed capacity (or supply) by exploiting differences in customer preference. Weekday opera performances are less popular than weekend performances. The Rockies know that a game against the Yankees will generate more demand than a game against the Cincinnati Reds. The highest electricity and telephone prices correspond to the periods of highest use and the lowest prices to the periods of lowest use.

Note that variable pricing is not pure group pricing, since it allows customers to self-select based on a menu of alternatives.

5.5.1 The Basics of Variable Pricing

Consider a hypothetical theme park that can serve up to 1,000 customers per day. The theme park charges a single admission price, and all rides are free after admission. During the summer, demand follows a stable and predictable pattern, with higher demand on weekends than during weekdays. We assume that demands for different days of the week are independent; that demand curves are linear, with intercepts, slopes, and satiating prices as shown in Table 5.3; and that the theme park has a marginal cost of zero per customer.

TABLE 5.3
*Intercepts and slopes for the price-response curves
in the theme park example*

Day of week	Intercept (D_i)	Slope (m_i)
Sunday	3,100	-62
Monday	1,500	-50
Tuesday	1,400	-40
Wednesday	1,510	-42
Thursday	2,000	-52.6
Friday	2,500	-55.6
Saturday	3,300	-60

What price should the theme park charge without variable pricing? Since its marginal cost is zero, it maximizes contribution by maximizing revenue. The revenue-maximizing single price can be found by solving the linear program

$$\max_{p,x} p \sum_{i=1}^7 x_i$$

pricing at some level of sophistication is used by the New York Mets, Anaheim Angels, New York Yankees, Atlanta Braves, Cleveland Indians, and San Francisco Giants in baseball and the Ottawa Senators and Pittsburgh Penguins of the National Hockey League. Popular variable-pricing techniques include higher prices for weekend games and for games against rivals and popular visiting teams.

Sporting teams like variable pricing because it can increase total revenue without increasing ticket prices. This is attractive in an environment where ticket price increases are likely to be condemned by fans and sportswriters as evidence of "greedy owners and players gouging loyal fans." In 2003 the New York Mets adopted a variable-pricing plan under which the price was increased \$10 for the 17 most popular games while lowering or holding prices unchanged for 43 games. "That allowed the Mets, who finished last in their division in 2002, to trumpet the plan as holding ticket price increases to only 4% on average" (Fatsis 2002).

While more and more teams are adopting some form of variable pricing, there are risks. Complex pricing schemes often draw criticism from fans. Furthermore, determining high-valued opponents before the season opens can be difficult. Some baseball teams have increased prices for games against the San Francisco Giants because there are a large number of fans willing to pay extra to see Barry Bonds play. However, an early season injury to Bonds would mean that these games would be suddenly overpriced and unlikely to sell out.

However, the potential gains are large enough that more and more clubs in more and more sports are adopting some form of variable pricing. The San Francisco Giants claim that their variable-pricing scheme netted them additional ticket revenue of \$1 million in 2002, a year in which their total profit was only \$100,000.

5.6.2 Passenger Airlines

Passenger airlines certainly meet the criteria for variable pricing—constrained capacity with demand that fluctuates over time. Southwest Airlines implements variable pricing directly by changing fares for their departures in order to match demand with the restricted capacity. However, the major "nondiscount" airlines also practice variable pricing. They do so by opening and closing the availability of various fare classes on different flights. Figure 5.2 shows average fares and load factors on a flight of a major airline from the Midwest to Florida during the winter. For operational reasons, the airline has scheduled the same aircraft on the flight each day. The airline is charging a higher average price on high-demand days (Saturday and Sunday) and lower prices on low-demand days (Tuesday, Wednesday, and Thursday). In this case, the airline is adjusting the average fare not by changing the list price but by opening and closing fare classes. When total demand is anticipated to be high, the airline will restrict availability of discount fares, causing the average fare to increase. When total demand is anticipated to be low, the airline will open availability of deep discounts, causing the average fare to drop. It is continually monitoring demand using its revenue management system to determine which classes to open and close on each flight. We describe how this is done in some detail in Chapters 6 and 7.

Discount airlines use variable pricing even more directly. If demand at Southwest Airlines begins to rise more quickly or more slowly than anticipated, Southwest may change

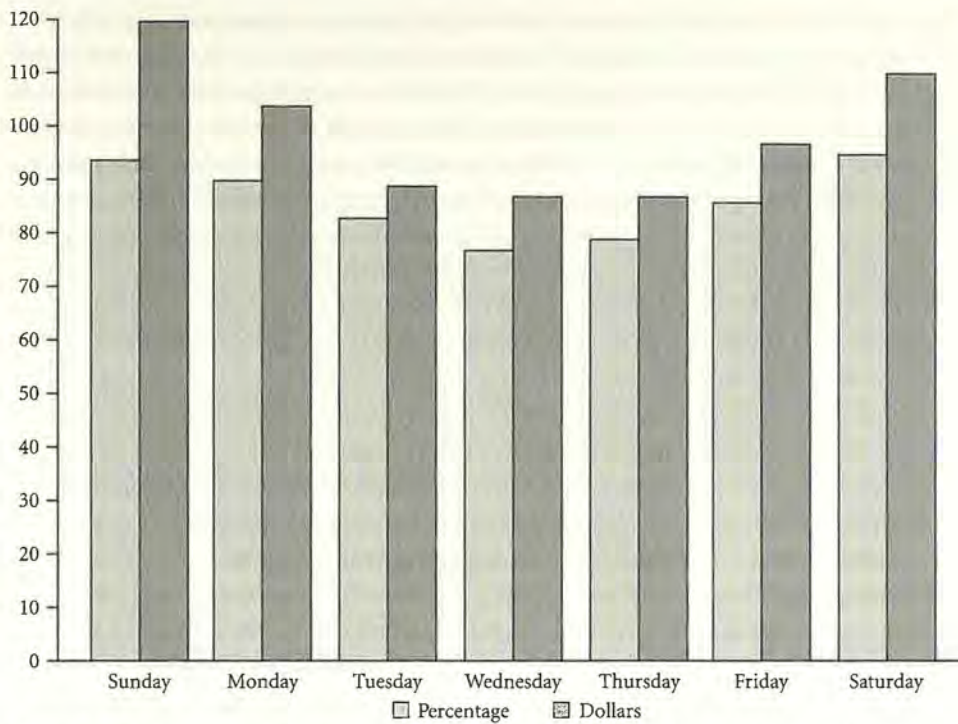


Figure 5.2 Average daily load factors for a Midwest-Florida flight on a domestic U.S. discount carrier. *Source:* Figure courtesy of Garrett J. van Ryzin, Professor, Columbia University.

the price. Southwest does not attempt to use prices to segment customers between “high-elasticity” and “low-elasticity” segments—rather, it is using variable pricing without segmentation to maximize the return from each flight and, to some extent, to smooth demand through demand diversion from high-fare/high-utilization flights to low-fare/low-utilization flights.

5.6.3 Electric Power

Electric power would seem to be an obvious candidate for variable pricing. Suppliers have constrained capacity, and electricity is difficult and expensive to store. Demand follows predictable patterns, and many customers have the flexibility and willingness to adjust their usage in response to prices. In fact, electric utilities have a further motivation to vary prices: There are tremendous cost differences between generation during peak and off-peak periods. For many utilities, it can be 10 times more expensive to generate power during a peak period than during an off-peak period. During a peak period, even a slight reduction in demand can lead to a substantial reduction in marginal production cost. In some cases, a 5% reduction in demand would reduce the marginal cost of production by 55%. This reduction would benefit all customers, not just those who responded to the variable price.⁵

Despite the obvious benefits that electric suppliers (and, potentially, consumers) could realize from time-of-day pricing, utilities in the United States have been slow to adopt

Revenue management (RM) refers to the strategy and tactics used by a number of industries—notably the passenger airlines—to manage the allocation of their capacity to different fare classes over time in order to maximize revenue. Revenue management is applicable under the following conditions.

1. The seller is selling a fixed stock of perishable capacity.
2. Customers book capacity prior to departure.
3. The seller manages a set of fare classes, each of which has a fixed price (at least in the short run).
4. The seller can change the availability of fare classes over time.

Revenue management can be considered a special case of pricing with constrained supply. However, the final two conditions give revenue management its special flavor. Revenue management is not based on setting and updating prices but on setting and updating the availability of fare classes, where each fare class has an associated fare (price) that remains constant through the booking period. This distinctive feature is a legacy from revenue management's origin. The passenger airlines that pioneered revenue management in the 1980s needed to utilize the capabilities they had at hand. This meant using the booking controls embedded in their reservation systems as the primary mechanism for controlling the fares displayed to customers at any time. Following the success at the airlines, revenue management has been adopted by a number of other industries, including hotels, rental cars, freight transportation, and cruise lines—many of whom use the same (or similar) reservation systems as the passenger airlines.

We use *revenue management industries* as a generic term for industries meeting the conditions listed earlier and *revenue management companies* to refer to companies that make use of booking controls and fare classes. While these companies use revenue management in its purest fashion, the techniques can be applied to any situation in which a seller needs to determine how to allocate a fixed supply among different channels or customer segments paying different prices.

This chapter serves as an introduction to revenue management. It is a prelude to the next three chapters, which address specific aspects of revenue management in more detail. We start by relating some of the history of revenue management and then outline revenue management strategy as an application of some of the price differentiation techniques introduced in Chapter 4. We then introduce revenue management tactics—the ways in which companies can manage their capacity to maximize return once they have chosen a strategy. Computerized reservation systems and global distribution systems play major roles in the way that revenue management is actually implemented. We describe their interaction and operation in Section 6.4. We describe the functions and operations of revenue management systems and how they operate in Section 6.6.2. Finally we discuss how the effectiveness of a revenue management program can be measured and cover the current status and future prospects of revenue management.

Since this chapter is introductory, it may feel a bit lacking in algorithms and computations. This will be made up in the following chapters, each of which addresses a specific aspect of tactical revenue management in detail: capacity allocation in Chapter 7, network management in Chapter 8, and overbooking in Chapter 9.

6.1 HISTORY

Prior to 1978, the airline industry in the United States was heavily regulated. Both schedules and fares were tightly controlled by the Civil Aeronautics Board (CAB). Fares were held sufficiently high to guarantee airlines a reasonable return on their investments. In 1978, Congress passed the Airline Deregulation Act. The act specified that the industry would be deregulated over four years, with complete elimination of restrictions on domestic routes and new service by December 31, 1981, and the removal of all fare regulation by January 1, 1983. The result was a shock from which the industry has not yet fully recovered.

One of the reasons Congress passed the Airline Deregulation Act was to inspire and encourage creative new entrants into the passenger airline business. One of the first new airlines to arise after deregulation was PeopleExpress. PeopleExpress was not unionized, and it offered a bare-bones service, with passengers paying extra for baggage handling and onboard meals. As a result, its cost structure was significantly lower than those of American and the other major airlines, such as United and Delta. By offering fares up to 70% below the majors, PeopleExpress filled its planes with passengers from a previously untapped market segment: price-sensitive students and middle-class leisure travelers who were induced to travel by the existence of fares far lower than anything the industry had seen before. PeopleExpress built its business initially by entering underserved markets, where its competition was primarily bus or car travel. PeopleExpress experienced four years of phenomenal growth and, in 1984, began service on Newark–Chicago and New Orleans–Los Angeles routes—key markets for American Airlines.

The choice American Airlines faced was stark. If it matched People's low fares, it could retain its customer base but would not be able to cover its costs. If it did not match People's fares, most of its customer base would be siphoned off by the low-price competitors. And, of course, in the deregulated environment, nothing could keep a successful PeopleExpress

from moving in on all of American's core markets. American Airlines seemed doomed—its only choice appeared to be between a slow death and a rapid one. However, Robert Crandall, CEO of American Airlines, formulated a counterattack: American would compete with PeopleExpress on low fares and simultaneously sell some of its seats at a higher fare as well.

In January 1985, American Airlines announced its "Ultimate Super Saver Fares" program. Ultimate Super Saver Fares matched PeopleExpress prices, with two key differences:

- For a passenger to qualify for an "Ultimate Super Saver" discount fare on American, he would need to book at least two weeks before departure and stay at his destination over a Saturday night. Passengers not meeting this restriction would be charged a higher fare. In contrast, PeopleExpress put no restrictions on their discount fares—every passenger paid a low fare.
- American restricted the number of discount seats sold on each flight in order to save seats for full-fare passengers who would be booking within the last two weeks prior to departure. PeopleExpress allowed every seat to be sold at a low fare.

American's two-pronged approach was carefully thought out and turned out to be the perfect counter to People's challenge. The booking restrictions ensured that the vast majority of the discount passengers buying American were leisure passengers who were able to book early and who were more price sensitive. The later-booking passengers who paid full fare were primarily business travelers who were less price sensitive but needed seat availability at the last minute. Both groups of passengers preferred American's superior service to People's bare-bones approach. In effect, American Airlines had segmented the market between leisure and business travelers and used differentiated pricing to attack a competitor.

The impact of American's actions was dramatic. American announced the new fares and new structure in January 1985. By March of that year, PeopleExpress was struggling and by August it was on the verge of bankruptcy. In September, Texas Air bought PeopleExpress for less than 10% of the market value it had enjoyed a year before. PeopleExpress' flamboyant CEO, Donald Burr, put the blame for his airline's demise squarely on American's superior yield management capability: "We had great people, tremendous value, terrific growth. We did a lot of things right. But we didn't get our hands around the yield management and automation issues."¹

This was the origin of revenue management (or *yield management* as it was called at the time). American's success prodded other major airlines in the United States to develop revenue management capabilities of their own. Throughout the remainder of the 1980s and well into the 1990s, carriers such as United, Delta, and Continental invested millions of dollars in implementing computerized revenue management systems and establishing revenue management organizations. As world aviation markets were increasingly deregulated, carriers in Europe and Asia began to adopt revenue management as well. Hotels and rental car companies followed the airlines in adopting revenue management. Marriott was a pioneer in hotel revenue management, and Hertz and National were pioneers in rental car revenue management. Vendors such as PROS and Talus Solutions developed and sold commercial revenue management software packages. The next wave of adopters included cruise lines,

passenger trains, and various modes of freight transportation. Development and investment in revenue management continues in many of these industries today.

6.2 LEVELS OF REVENUE MANAGEMENT

Successful revenue management requires consistent execution at three levels, as shown in Table 6.1. *Revenue management strategy* is the identification of customer segments and the establishment of products and prices targeted at those segments. Once products and prices have been established, *revenue management tactics* require setting and updating limits on how much of a particular product can be sold at a particular fare to each segment for some period of time—say, a day or a week. *Booking control* is the moment-to-moment determination of which booking requests should be accepted and which should be rejected.

TABLE 6.1
The three levels of revenue management decisions

Level	Description	Frequency
Strategic	Segment market and differentiate prices	Quarterly or annually
Tactical	Calculate and update booking limits	Daily or weekly
Booking control	Determine which bookings to accept and which to reject	Real time

In most cases, booking control is a simple, mechanical process of checking whether a booking can be accepted given the booking limits currently in place. For airlines, hotels, rental car companies, and cruise lines, booking control is a function of the reservation system. Tactical revenue management is the “brains” of the process. It is where future demand is forecast, optimization algorithms run, and booking limits set and updated.

We first discuss the strategy of revenue management and how it applies in different industries. We then turn to a somewhat detailed discussion of booking control—the “real-time” face of revenue management. With this background in place, the next three chapters deal with elements of tactical revenue management—capacity control, network management, and overbooking.

6.3 REVENUE MANAGEMENT STRATEGY

Revenue management strategy consists of the identification of customer segments and the establishment of products targeted at those segments. A fundamental element of revenue management strategy at many hotels, rental car companies, and airlines is the distinction between leisure customers and business customers first recognized by American Airlines. The different characteristics of leisure and business customers are shown in Table 6.2. The airlines used these characteristics to segment their market and create virtual products oriented toward the different segments. A typical approach is shown in Table 6.3. Here, an airline has identified five customer segments, two business and three leisure, with at least one product targeted toward each segment. The leisure products have various restrictions (early purchase, Saturday night stay) that make them unattractive or unavailable to many business

TABLE 6.2
Characteristics of business vs. leisure customers at the passenger airlines

Leisure	Business
Highly price sensitive	Less price sensitive
Book earlier	Book later
More flexible to departure and arrival times	Less flexible
More accepting of restrictions such as Saturday night stayovers	Less accepting of restrictions

TABLE 6.3
Typical airline market segmentation

	SEGMENT					
	BUSINESS ORIENTED			LEISURE ORIENTED		
Price sensitivity	Low	Moderate	Sensitive	Moderate	Very high	
Schedule flexibility	Need	Important	Somewhat	Prefers	Very flexible	
Inventory	Last seat	Somewhat restricted	Restricted	Restricted	Very limited	
Product	Unrestricted Business	Corporate Discount	Discount Business	Regular Leisure	Sale fares Web only Priceline.com	

travelers. In the terminology of Section 4.3, the airlines used artificial restrictions to create an inferior product that they could sell at a lower price to a more price-sensitive segment of the market.

Product versioning is not the only price differentiation tactic used by the airlines. They also use most of the other tactics described in Section 4.3 in some form or other. Airlines sell products targeted to many segments such as government, senior citizens, groups, tour operators, and cruise lines. International airlines are great believers in regional pricing. They will sell the same tickets for different prices in different countries (even after adjusting for exchange rates) in order to exploit differences in price sensitivity. Airlines are also enthusiastic channel pricers; ticket prices on the Internet are often cheaper than through travel agents.

Other revenue management companies followed the airlines in segmenting their market, creating virtual products and establishing a wide range of prices. Hotels and rental car companies establish and manage products and different prices oriented toward various groups: corporate, leisure, and business segments. As a simple example, many hotels in Hawaii maintain so-called *kamaaina* rates for local residents.² These rates are often far lower than the standard rates, and hotel revenue managers make them available during slack periods in order to fill their properties. Only customers who can prove Hawaiian residency qualify for a *kamaaina* rate. This is a classic example of group pricing. By confining availability to locals, hotels ensure that the low *kamaaina* rates do not cannibalize their high-paying business from the mainland United States and Asia.

Cruise lines do not have late-booking business customers, but they do sell to the *incentive* segment—bulk purchases of cruise capacity by companies to be used as incentives for their employees. For example, a medical supply company might purchase 20 berths on a particular sailing to reward the 20 top North American sales managers. This incentive business is

booked earlier and is sold at lower fares than individual cruises. Cruise lines also charge different rates for customers from different cities. Managing availability by city is one way they can maximize contribution from each sailing.³

6.4 THE SYSTEM CONTEXT

American Airlines was able to outmaneuver PeopleExpress in part because it had a developed a *computerized reservation system* called SABRE that allowed it to save seats for later-booking business passengers—a capability PeopleExpress lacked. When it was turned on in 1964, SABRE was a technological marvel. Not only did it replace the colored index cards in revolving trays that American and other airlines had used since the 1930s to manage their reservations, it gave American Airlines the ability to distribute its products and fares globally and to receive bookings from anywhere in the world.

SABRE combines the functions of a *computerized reservation system* (CRS) and a *global distribution system* (GDS). The key function of a reservation system is to serve as the repository of all the bookings that have been accepted for future flights. The reservation system also contains the controls that specify how many bookings from different fare classes the airline will accept on future flight departures. A reservation system is a critical capability for any revenue management company. Many large airlines, hotels, and rental car companies have developed their own systems, while many smaller companies are “hosted” by commercial providers of reservation services.

Figure 6.1 illustrates some of the distribution channels for a typical airline. At one time, more than 80% of the bookings for a major airline would be communicated through a global distribution system. Several airlines owned their own systems: American Airlines

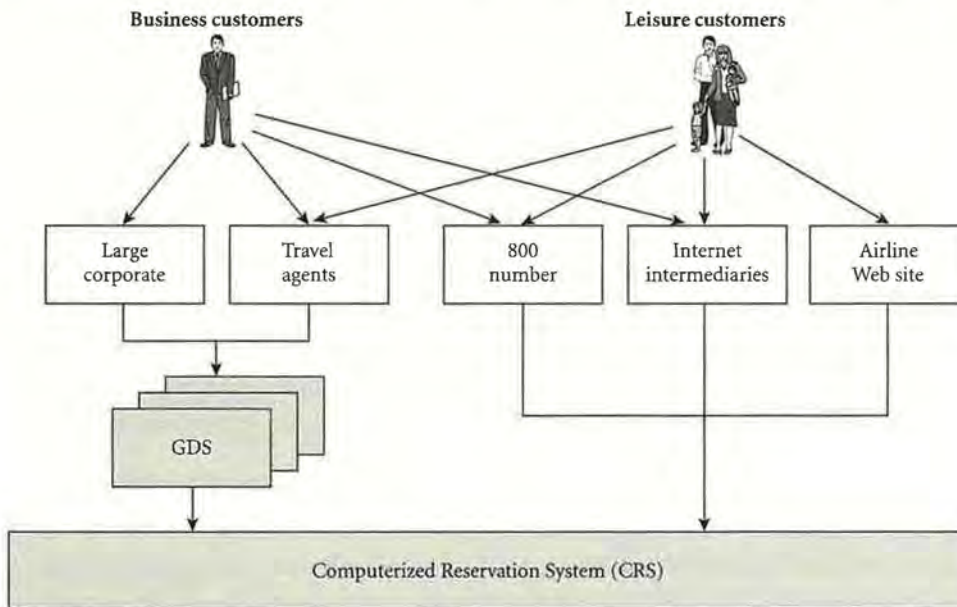


Figure 6.1 Distribution channels for a typical airline.

owned SABRE, United Airlines owned Galileo (then called Apollo), and Northwest Airlines owned Worldspan. Other global distribution systems, such as ABACUS, were owned by consortia of airlines. However, following a series of Justice Department rulings that eliminated the ability of an airline to give preferential treatment to its own flights on its GDS, airlines largely divested themselves of their global distribution systems.

Global distribution systems provide a conduit connecting the reservation system and a number of important sales outlets—particularly travel agents. There are currently over 180,000 GDS terminals worldwide. Most of these terminals are placed with travel agents, but some are in the travel offices of large companies (Marvel 2004). The four most important global distribution systems are AMADEUS, Galileo, SABRE, and Worldspan.⁴ Each GDS distributes schedule, fare, and booking availabilities to the agencies and transmits bookings and cancellations from the agencies to the airlines. Most airlines, major hotel chains, major rental car companies, and cruise lines have contracts with most or all of the global distribution systems allowing them to distribute and sell their products. There are exceptions: Both Southwest Airlines and JetBlue Airways subscribe only to SABRE and do not accept bookings through any other GDS (Kontzer 2004).

This system background is important because airlines did not choose to pursue a revenue management strategy and then design reservation systems to support that strategy. Rather, reservation systems were developed first and revenue management capabilities were added later. SABRE went operational in 1964—14 years before the airlines were deregulated and 21 years before American first did battle with PeopleExpress. The computerized reservation systems were among the most sophisticated and complex transaction-processing systems of their day. But they were not very flexible. Since they were built on large mainframes using 1960s-vintage software, major modifications were (and are) expensive and time consuming. As a result, the way airlines and other travel companies manage their inventory is still dictated by the capabilities of these systems. And at the heart of these systems are the booking control mechanisms that determine moment by moment which fare classes are available for sale and which are not.

6.5 BOOKING CONTROL

Booking control is the real-time face of revenue management. The function of booking control is to determine whether or not each booking request received should be accepted or rejected.⁵ As shown in Figure 6.1, booking requests stream into the reservation system from many different channels. Within a very short time (typically less than 200 milliseconds), the reservation system needs to send a message whether or not that request can be accepted. Given the short time available, reservation systems typically rely on very simple, mechanical procedures for determining whether or not to accept a request.

When a booking request is received—say, for a seat for a future flight or a hotel room for some future dates—the request is assigned to a fare class. This assignment may be based on the time and characteristics of the request (e.g., whether it meets the qualifications for a deep discount), the channel through which the request was received, the market segment of the customer (e.g., group versus individual), or combinations of all of these. The fare classes are typically assigned letters, so we can speak of a B-Class or M-Class booking request.⁶

The reservation system includes a *booking limit* for each fare class on each product. When a booking request is received, the reservation system checks the booking limit for the associated fare class. If there is sufficient availability, the request will be accepted; if not, the request will be rejected.

Example 6.1

An airline receives a B-Class request for 3 seats on Flight 137 from Houston to Miami, departing in two weeks. The current B-Class booking limit for this flight is 2. Because there is insufficient availability, the request is rejected.

At each time, the booking limits currently in place will determine which booking requests will be accepted and which rejected. When a new booking is accepted or a previously accepted booking cancels, the reservation system automatically updates the booking limits for that product. This is all done very rapidly, consistent with the need of the reservation systems to work in a real-time environment.

6.5.1 Allotments

An obvious way to manage bookings would be to divide the available capacity into discrete chunks and to allocate each chunk to a fare class. This is known as the *allotments* approach, and the size of the chunk allocated to each fare class is called its allotment. Under this approach, bookings are accepted in a class until the allotment for that class is exhausted.

Example 6.2

A 100-seat aircraft is being managed using allotments. Thirty seats have been allotted to deep-discount bookings (B-Class) with a \$125 fare, 45 seats to full-fare coach (M-Class) with a \$200 fare, and 25 seats to business class (Y-Class) with a \$560 fare. Two weeks before departure, 25 B-Class bookings, 45 M-Class bookings, and 10 Y-Class bookings have been accepted. The remaining allotments are 5 seats for B-Class, no seats for M-Class, and 15 seats for Y-Class.

While the allotments approach is easy to understand, it has a major drawback: It doesn't work very well. In particular, it can result in high-fare customers being rejected while lower fare customers are still being accepted. In Example 6.2, once the M-Class allotment closed, the airline would be accepting customers paying \$125 while rejecting \$200 bookings. This is a prime revenue management sin—we cannot maximize revenue by rejecting high-fare customers in order to save seats for low-fare customers. For this reason, most revenue management companies “nest” their inventory so that high-fare customers have access to all of the inventory available to lower-fare customers.

6.5.2 Nesting

Nesting was developed to avoid the situation in which high-fare bookings were rejected in favor of low-fare bookings. To describe nesting, we number fare classes so that 1 is the

highest fare class and n is the lowest. We define b_i as the booking limit for class i . With nested booking controls, booking limits are always nondecreasing; that is,

$$b_1 \geq b_2 \geq b_3 \geq \dots \geq b_n \quad (6.1)$$

Furthermore, at each time, every fare class has access to all of the inventory available to lower fare classes. This avoids the possibility of accepting low-fare bookings while rejecting high-fare bookings.⁷ The booking limit for the highest class, b_1 , is an upper bound on the total number of bookings that will be accepted. If the airline is not expecting any no-shows or cancellations, then b_1 would equal the capacity of the flight. When there is a possibility that bookings will cancel or not show, then it may be optimal to *overbook*, that is, to accept more bookings than available capacity.⁸

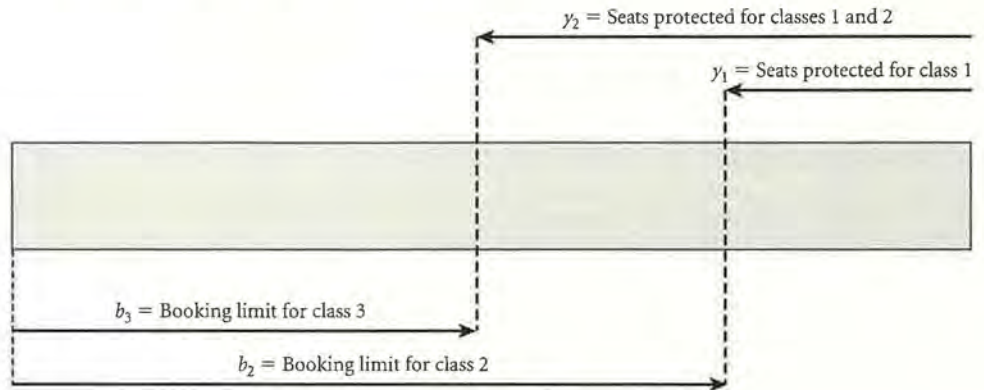


Figure 6.2 Nested booking limits and protection levels.

We can also describe nesting in terms of *protection levels*. The protection level for class i is the total number of seats available to class i and all higher classes. Let y_j be the protection level for class $j = 1, 2, 3, \dots, n-1$. The relationship among booking limits and protection levels with nested booking controls is illustrated in Figure 6.2. From this figure it should be evident that

$$y_j = b_1 - b_{j+1} \quad \text{for } j = 1, 2, \dots, n-1 \quad (6.2)$$

and $y_n = b_1$; that is, the protection level for the lowest fare class is equal to the total booking limit for the flight. Combining Equation 6.2 with Condition 6.1, we can see that protection levels decrease for higher fare classes; that is,

$$0 \leq y_1 \leq y_2 \leq \dots \leq y_{n-1}$$

as can be seen in Figure 6.2.

6.5.3 Dynamic Nested Booking Control

What happens when we accept a booking when limits are nested? One approach is to decrement all nonzero booking limits by 1 every time we book a seat. Once a booking limit for a class reaches zero it remains there, at least until the next reoptimization.

Example 6.3

A flight has a total booking limit of 100 and five fare classes with booking limits of $(b_1, b_2, b_3, b_4, b_5) = (100, 73, 12, 4, 0)$. Note that this flight is currently not accepting any bookings from class 5. Any booking class with a limit of 0 is said to be *closed*. We can derive the corresponding protection levels using Equation 6.2: $(y_1, y_2, y_3, y_4, y_5) = (27, 88, 96, 100, 100)$.

Table 6.4 shows a series of booking requests, whether or not each request would be accepted, and the updated booking limits and protection levels, starting from the booking limits specified in Example 6.3. Each line of the table lists the booking limits and protection levels just prior to receiving the request. The last column shows the corresponding action for each request.⁹

TABLE 6.4
Dynamics of booking limits and protection levels

	BOOKING LIMITS					PROTECTION LEVELS					Request	Action
	1	2	3	4	5	1	2	3	4	5		
1	100	73	12	4	0	27	88	96	100	100	2 seats in Class 5	Reject
2	100	73	12	4	0	27	88	96	100	100	5 seats in Class 2	Accept
3	95	68	7	0	0	27	88	95	95	95	1 seat in Class 2	Accept
4	94	67	6	0	0	27	88	94	94	94	1 seat in Class 4	Reject
5	94	67	6	0	0	27	88	94	94	94	3 seats in Class 3	Accept
6	91	64	3	0	0	27	88	91	91	91	4 seats in Class 3	Reject
7	91	64	3	0	0	27	88	91	91	91	2 seats in Class 3	Accept
8	89	62	1	0	0	27	88	89	89	89	4 seats in Class 3	Reject
9	89	62	1	0	0	27	88	89	89	89	1 seat in Class 3	Accept
10	88	61	0	0	0	27	88	88	88	88	2 seats in Class 2	Accept
11	86	59	0	0	0	27	86	86	86	88	2 seats in Class 3	Reject

Note that the protection levels for the higher classes are preserved while bookings are being accepted for the lower classes. Even after 14 total bookings have been accepted, 27 seats are still protected for Class 1 demand. Nesting implies that a class can be closed as a result of accepting bookings for a higher class. In Table 6.4, this happens with the second booking request received—accepting five Class 2 bookings led to the closure of Class 4. Finally, nesting guarantees that fare classes close in order, from lowest to highest.¹⁰

6.5.4 *Managing Cancellations

Since almost 30% of airline bookings are cancelled before departure, the bookings management process at a passenger airline needs to be able update booking limits for cancellations. The obvious way to do this would be to treat a cancellation as the opposite of a booking; that is, when a single-seat booking cancels, increase all the positive booking limits by 1, just as we decreased all the positive booking limits by 1 when we accepted the booking. This is, in fact, the way that many reservation systems treat cancellations.

There is a problem with this approach, however. Let's say we have a booking limit of three seats in the lowest booking class and we accept a booking for four seats in some higher class. Then, the lowest class would close. Now, assume that 10 seconds later the booking for four

seats is cancelled (perhaps the original request was an error). Under the approach described earlier, the four seats would be added back to the open booking classes and the lowest booking class would remain closed. This seems wrong. After all, we have the same available capacity and the same forecasts for future demand that we had 10 seconds before—why should our booking limits change?

Another way of viewing the issue is that, under the process we have so far described, the closure of a booking class is irreversible, at least until reoptimization occurs.¹¹ For this reason, we call the procedure illustrated in Table 6.4 an *irreversible* process. Recognizing the shortcomings of an irreversible process, some reservation systems enable a *reversible* bookings management process. The trick in creating a reversible process is to allow the booking limits to go below zero. In a reversible process, new bookings are subtracted from all booking limits and cancellations are added back to all booking limits. As before, a booking request is accepted only if the corresponding booking limit is greater than the number of seats requested. This means that a class with a negative booking limit is closed. With a reversible process, a cancellation can then result in availability going from zero (or less than zero) to positive, opening a previously closed class.

These two approaches to treating cancellations are illustrated in Table 6.5, which shows a sequence of booking requests and cancellations and the corresponding evolution of the booking limits for four booking classes under both approaches. The initial set of booking limits is (73, 12, 4, 0). The second entry in each row of Table 6.5 is either a booking request or a cancellation. The other entries in each row are the booking limits just prior to the event under both approaches and the action (if any) that the airline would take. Since an airline cannot “reject” a cancellation, there are no actions associated with cancellations.

From Table 6.5 we can see that the reversible approach sometimes opens previously closed booking classes. This means that it accepts more low-fare bookings than the irreversible approach. While this may seem appealing, the other side of the coin is that the irreversible approach saves more seats for later high-fare bookings. Generally, it is felt that the reversible approach provides higher revenue, since a booking request that immediately cancels will not

TABLE 6.5
Dynamics of booking limits and protection levels

		BOOKING LIMITS					BOOKING LIMITS				
Event		1	2	3	4	Action	1	2	3	4	Action
1	Request 2 seats in Class 3	73	12	4	0	Accept	73	12	4	0	Accept
2	Request 2 seats in Class 3	71	10	2	0	Accept	71	10	2	-2	Accept
3	2 bookings cancel	69	8	0	0		69	8	0	-4	
4	Request 5 seats in Class 3	71	10	0	0	Reject	71	10	2	-2	Reject
2	Request 3 seats in Class 2	71	10	0	0	Accept	71	10	2	-2	Accept
3	Request 2 seats in Class 3	68	7	0	0	Reject	68	7	-1	-5	Accept
4	3 bookings cancel	68	7	0	0		68	7	-1	-5	
5	Request 2 seats in Class 3	71	10	0	0	Reject	71	10	2	-2	Accept
6	Request 2 seats in Class 3	71	10	0	0	Reject	69	8	0	-4	Reject
7	2 bookings cancel	71	10	0	0		69	8	0	-4	
8	Request 1 seat in Class 2	73	12	0	0	Accept	71	10	2	-2	Accept
9	Request 1 seat in Class 3	72	11	0	0	Reject	70	9	1	-3	Accept
	Final state	72	11	0	0		69	8	0	-4	

change the booking limits. (For an example of this, see event numbers 2 and 3 in Table 6.5 and the resulting changes in booking limits under the two approaches.) However, which approach is used is often determined by the capabilities of its reservation system rather than by considerations of relative benefit.¹²

6.6 TACTICAL REVENUE MANAGEMENT

The job of tactical revenue management is to calculate and periodically update booking limits. These booking limits are transmitted to the reservation system, which then uses them to determine which booking requests to accept and which to reject based on the logic described in Section 6.5.

To introduce the tactical revenue management problem in its most general form, it is useful to define *resources*, *products*, and *fare classes*.

- *Resources* are units of capacity managed by a supplier. Examples of resources are a flight departure, a hotel room night, and a rental car day. Each resource is constrained—the flight departure has a limited number of seats, the hotel only has a limited number of rooms it can sell, and a rental car company has only so many cars it can rent out on a particular day.
- *Products* are what customers seek to purchase. A product may require use of one or more resources. A seat on Flight 130 from St. Louis to Cleveland on Monday, June 30, is a product that uses only a single resource. A two-night stay at the Sheraton Cleveland for a customer arriving on March 19 and departing on March 21 is a product that uses two resources: a room night on March 19 and a room night on March 20.
- There are one or more *fare classes* associated with each product. Each fare class is a combination of a price and a set of restrictions on who can purchase the product and when. As discussed in Section 6.3, different fare classes can be used to establish different virtual products, for group pricing, for regional pricing, or for combinations of all of these.

Table 6.6 shows how the resource/product approach applies in four different revenue management industries. A *resource unit* is the smallest unit of capacity that can be sold: a single seat for an airline, a room night for a hotel. A company may sell several different *types* of resource: airlines sell coach, business, and first-class seats; rental car companies offer subcompact, compact, mid-size, and luxury cars; and a hotel may sell five or six different room types. While resources are the assets a company needs to manage, products are what customers actually want to *buy*. An airline's products are direct and connecting flight combinations. A hotel's products are combinations of arrival date and length of stay. A customer reserving a two-night stay starting next Wednesday is buying a different product than one reserving a two-night stay starting next Thursday.

The resource/product/fare class terminology enables us to formulate the tactical revenue management problem in a very general manner.

TABLE 6.6
Revenue management in four different industries

	Passenger airlines	Hotels	Rental car companies	Cruise lines
Resource	Flight leg	Property	Location	Sailing
Resource unit	Seat	Room night	Rental day	Berth
Resource types	Service classes	Room types	Car types	Berth classes
Number of types	1–3	1–many	3–10	5–10
Products	Itineraries	Arrival date/ length of stay	Arrival date/ length of rental	Cruise
Products per resource	Many for hub-and-spoke airline	Many	Many	Few

A supplier controls (1) a set of resources with fixed and perishable capacity, (2) a portfolio of products consisting of combinations of one or more of the resources, and (3) a set of fare classes associated with each of the products. The tactical revenue management problem is to choose which fare classes should be open and which closed for sale at each moment in order to maximize expected total net contribution.

The fact that revenue management is based on opening and closing fare classes does not make a big difference from the customer's point of view—customers just see the lowest available fare changing over time. The focus on opening and closing fare classes is partly due to the design of the reservation systems. It is also based on the competitive dynamics of the airline market. Specifically, most airlines believe they need to match the advertised, or “headline,” fares offered by their competition in key markets in order to drive demand for bookings. However, as bookings arrive, revenue management enables them to “shape” demand to the limited capacity of each flight. Revenue management supplements rather than replaces pricing.

6.6.1 Components of Tactical Revenue Management

As we have seen, the point of tactical revenue management is to determine which fare classes should be open and which closed for all products in order to maximize return from a fixed set of resources. This general problem can be decomposed into three constituent problems.

1. *Capacity allocation.* How many customers from different fare classes should be allowed to book?
2. *Network management.* How should bookings be managed across a network of resources, such as an airline hub-and-spoke system or multiple-night hotel stays?
3. *Overbooking.* How many *total* bookings should be accepted for a product in the face of uncertain future no-shows and cancellations?

Table 6.7 gives a general view of the relative importance of each of these problems in different industries.

Capacity allocation is important whenever a company sells the same unit of constrained

TABLE 6.7
Relative importance of revenue management problems among industries

Industry	Capacity allocation	Network management	Overbooking
Passenger airlines	*****	****	*****
Hub-and-spoke	****		*****
Point-to-point	***	N/A	*****
Hotels			
Business	*****	*****	****
Resort	****	****	***
Rental cars	*****	*****	*****
Cruise lines	*****	N/A	***
Passenger rail	****	*****	***
Air freight	*****	****	*****
Event ticketing	*****	N/A	N/A

* = little importance, ** = can be important in some cases, *** = usually important, **** = of great importance, N/A = not applicable.

capacity or inventory at two or more different prices. In this case, the company needs to determine how many units to sell at the lower price(s) and how many to reserve for sale at a higher price. The major passenger airlines need to determine how many seats to sell to early-booking groups and discount customers and how many to reserve for later-booking business customers. Rental cars and hotels face similar challenges. Sporting teams need to determine how many tickets to sell on discounts or promotions versus holding to sell at full price. Air freight carriers have agreements with different customers that specify different rates—the carriers want to make sure that high-paying customers do not find themselves shut out because all capacity was reserved by lower-paying customers. On the other hand, resort hotels, discount carriers such as Southwest, and passenger trains may only maintain a small number of prices for the same inventory. For these companies, capacity allocation tends to be less important than other aspects of revenue management.

Network management is important for companies that sell products consisting of combinations of resources. It is the single most important component of revenue management for passenger railways and for rental cars and business hotels, where managing length of stay has more impact on total revenue than managing rate classes. It is important for airlines that operate hub-and-spoke systems but not as important as capacity allocation or overbooking. It is of little or no importance in industries in which each product sold uses only a single resource, such as cruise lines, point-to-point airlines, sporting events, and theater.

Overbooking is important whenever bookings are allowed to cancel or not show with little or no penalty. Since airline tickets are typically sold in this way, overbooking has been a major component of airline systems. It has become somewhat less important as the fraction of nonrefundable tickets has increased, but it is still a very important component of revenue management at most airlines, as it is at business hotels and rental car companies. On the other hand, most resort hotels and cruise lines do not overbook because of the high perceived cost of denying service to a customer who may have planned his entire vacation (or even honeymoon) around staying at a particular resort or departing on a particular cruise. It is generally not used for sporting events and theater, where tickets are nonrefundable and seats are individually assigned.

6.6.2 Revenue Management Systems

The primary job of a revenue management system is to calculate and update the booking limits within the reservation system. Typically a revenue management system is separate from the reservation system but linked to it, as shown in Figure 6.3. The revenue management system receives a feed of bookings and cancellations from the reservation system and, in turn, calculates booking limits that are transmitted periodically to the reservation system. The revenue management system will then send messages to update the booking limits in the global distribution systems.

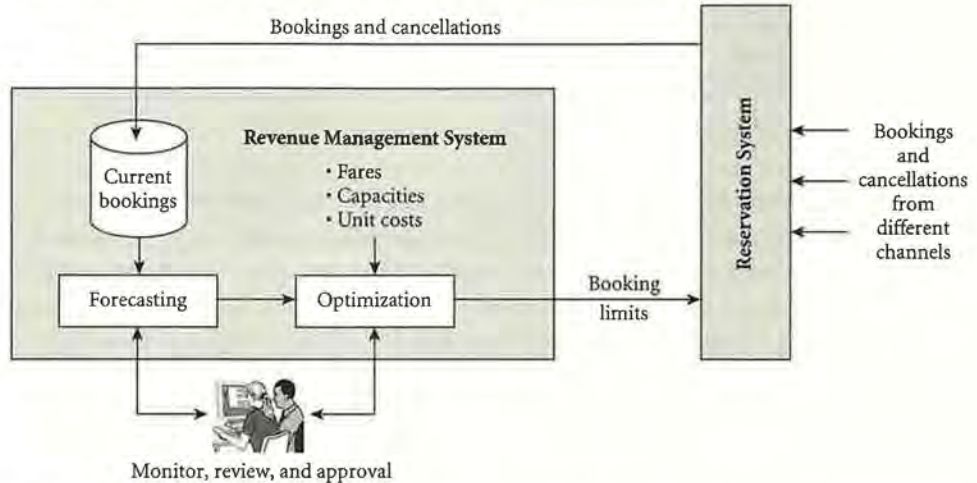


Figure 6.3 Schematic overview of a typical revenue management system.

The revenue management system includes a database with information on current bookings on all flights. It also includes a database incorporating fares for all product/fare class combinations, capacities on all flight legs, and passenger variable costs by product. This information is typically extracted from other computer systems, specifically the pricing system, the scheduling system, and various accounting systems.

The *forecasting* module generates and updates forecasts for all product/fare class combinations for all future dates. Generally, an initial forecast will be generated about one year prior to departure—the time when most airlines first allow bookings for a future flight. This forecast will be updated periodically as bookings and cancellations are received over time. Typically, a forecast will be updated monthly when a flight is six months or more from departure and more frequently as departure approaches. Forecasts are typically updated daily for flights that are within two weeks of departure.

The forecasts generated by the revenue management system are *probabilistic*; that is, they predict both a mean and a standard deviation for future demand. Uncertainty plays a major role in the calculation of optimal booking limits.

Example 6.4

A typical forecast generated by a revenue management system would be that expected future B-Class bookings on Flight 47 from Chicago to Grand Rapids,

Michigan, departing at 9:30 two weeks from today is 13 passengers with a standard deviation of 9, while expected M-Class bookings will be 22 passengers with a standard deviation of 10.

These forecasts are based on current and recent bookings for this flight, historic booking performance for this flight, and general demand trends for all flights. Most systems also forecast cancellations and no-shows in order to support overbooking calculations.¹³ Standard forecasting techniques are used to generate the initial forecast and to update the forecasts as bookings and cancellations are received over time. Chapter 9 of Talluri and van Ryzin (2004) describes some of the leading forecasting techniques used to support revenue management. Makridakis, Wheelwright, and Hyndman (1997) is a good general introduction to forecasting.

The probabilistic forecasts are the primary input used by the optimization module to generate the booking limits. As described in Chapters 7 and 8, booking limits are calculated by estimating the economic tradeoff between accepting more discount bookings now versus having more capacity available to serve future bookings. Typically, booking limits are automatically recalculated whenever the demand forecasts are updated.

Revenue management systems recommend booking limits for each product/fare class combination. Effective revenue management requires that these recommendations be continually monitored and reviewed. For example, Delta Airlines employs a group of more than 30 “revenue management analysts” who continually monitor the forecasts and booking limits generated by Delta’s revenue management system. Monitoring and adjusting forecasts is a particularly important part of the process. For example, the revenue management system will not know that the Super Bowl will be held in New Orleans this year. It is the job of the revenue management analyst to adjust the forecasts for January flights in the New Orleans market to reflect the additional demand generated by the Super Bowl.

6.6.3 Updating Booking Limits

Most RM companies set initial booking limits based on resource capacities, the demand forecasts for each booking class, and the economic tradeoffs among the classes. Most reservation systems begin to accept bookings about a year prior to a flight departure or room rental date. At this time an airline sets initial booking limits for fare classes on those departures. The initial booking limits are loaded into the reservation system and decremented as bookings are accepted. Periodically, demand is reforecast and new booking limits are calculated based on the new forecast of demand and remaining unbooked capacity. This is known as a *reoptimization* or an *update*. Updates can be triggered in three different ways.

1. *Periodic updates* occur at scheduled intervals. These intervals are typically long (say, monthly) when the flight is far from departure and booking requests are rare. As the flight approaches departure and the bookings pace increases, periodic updates are scheduled more frequently—daily or even more often during the last two weeks prior to departure.

2. *Event-driven updates* are triggered by events such as a booking class closing, a change in aircraft, and an unanticipated spike in demand.
3. *Requested updates* may be launched at any time by a flight controller or revenue manager based on competitive actions, changes in fares, anticipated changes in future demand, or any other reason.

Each time booking limits are updated for a flight, the new booking limits are loaded into the reservation system and the booking management process resumes immediately starting from the new limits.

6.7 NET CONTRIBUTION IN REVENUE MANAGEMENT

Revenue management decisions—like other pricing and revenue optimization decisions—are usually based on maximizing *expected net contribution*. The expected net contribution of a customer commitment is the price (fare) minus the incremental cost of the commitment. When passenger airlines in the United States first began to adopt revenue management, they were in the happy situation that their fares were well above their incremental costs (or *passenger variable costs*, as they are called in the airline industry). The earliest generation of airline revenue management systems ignored incremental costs (or assumed they were zero) and focused on maximizing expected revenue (hence the name *revenue management*). As a result, much of the revenue management literature is couched in terms of maximizing revenue instead of maximizing contribution.

As long as revenue management was confined to the passenger airlines and fares remained high relative to incremental cost, this didn't make much difference; maximizing expected revenue led to virtually the same decisions as maximizing expected net contribution. However, in the 1990s, two things began to happen. First, discount fares began a steady decline, to the point where by 2001, the difference between the deepest discount fares and incremental cost was small. Second, revenue management began to be adopted by new industries. In some industries, such as freight transportation and cruise lines, the incremental cost associated with a customer commitment could be significant. Contribution in these industries could be maximized only if incremental costs were incorporated explicitly into revenue management decisions.

6.7.1 Computing Incremental Costs

Table 6.8 shows some of the elements of incremental cost and a subjective estimate of the importance of incremental cost in several revenue management industries. As noted, incremental costs are relatively high (and thus important) in cruise lines and container shipping. Food is a major expense for cruise lines, and the amount of food they purchase and load for a sailing depends on the number of passengers booked. At the other end of the spectrum, incremental costs are relatively low (and therefore generally unimportant) in theater and sporting events. It generally costs little more to stage a play in front of a full theater versus an empty one—not counting, of course, the psychic cost to the actors of playing before an empty house. In between these two extremes fall hotels, rental car companies, and

TABLE 6.8
Incremental cost elements in some revenue management industries

Industry	Incremental cost elements	Relative importance
Passenger airlines	Commissions, incremental fuel, meal costs, passenger fees	Moderate
Hotels	Commissions, room cleaning, wear and tear	Low
Rental cars	Commissions, check-in/check-out, wear and tear	Moderate
Cruise lines	Commissions, food, cleaning costs	High
Container shipping	Repositioning, handling, incremental fuel	High
Sporting events and theater	Commissions	Very low

passenger airlines. In all cases it becomes more important to estimate incremental costs accurately and to incorporate them into revenue management decisions when prices are low.

Incremental costs may vary by channel and market segment. In particular, commissions and distribution fees are important components of incremental cost in many revenue management industries. SABRE charges \$3 to \$4 per flight segment for an airline booking and about \$4 per room for a typical hotel booking. In addition, travel agents receive a commission of about 10% for every booking they make. On the other hand, the variable cost of a direct reservation on a hotel's own Web site is about \$1.50 (Marvel 2004). Figure 6.4 shows distribution costs for a \$325 ticket sold through different channels for a major U.S. airline in 2002. The cost per booking varies from less than \$10 for an opaque ticket to almost \$45 for a booking sold through a travel agency.¹⁴

The range of costs shown in Figure 6.4 represents a channel-pricing opportunity. It also provides an expanded opportunity for revenue management—airlines and hotels can establish fare classes based on the channel of a booking request and determine whether or not

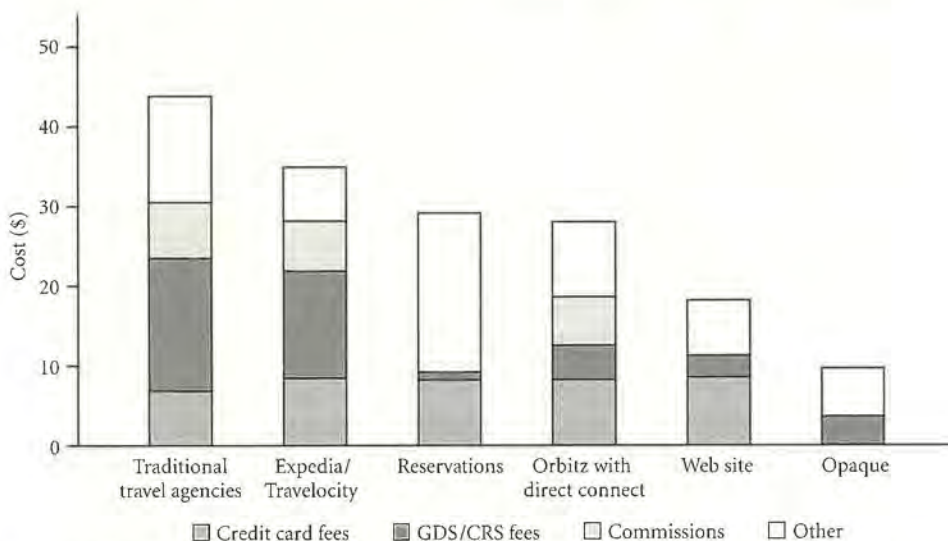


Figure 6.4 Per-ticket distribution costs by channel for a \$325 ticket for a major U.S. airline (2002). Source: Figure courtesy of Continental Airlines.

to provide availability through that channel. Furthermore, it provides a tremendous economic incentive for revenue management companies to steer demand as much as possible through lower-cost channels. While online travel has been booming, the global distribution systems still account for about 60% of bookings for a typical airline and about 68% for a typical hotel, leaving plenty of scope for future savings (Grossman 2004).

Calculation of incremental costs for a particular booking request requires an activity-based incremental costing model. A typical activity-based costing model for a fictional rental car company (Rent-a-Lemon) is shown in Table 6.9. Three types of cost element are represented: a transaction cost that is generated by every booking, a daily cost that depends on the number of days in the rental, and a commission. The commission is 15% for bookings through travel agents and 0 for bookings through Rent-a-Lemon's 800 number or at one of its airport counters.

TABLE 6.9
Incremental cost elements for Rent-a-Lemon

Cost element	Units	Value	Description
Commissions	\$/	15% 0%	Commission paid for bookings through travel agents Commission paid for bookings through other channels
Transaction cost	\$/Rental	\$8.25	Paperwork, cleaning costs, check-in and check-out costs for each rental
Daily cost	\$/Day	\$4.32	Wear and tear per day of rental

Example 6.5

Both Mr. Smith and Mrs. Jones want to rent a car for two days from the Rent-a-Lemon location at the Topeka airport. The rate available to both Smith and Jones is \$28.00/day. Mr. Smith is making his request at the Rent-a-Lemon counter at the Topeka airport, while Mrs. Jones is requesting the rental through a travel agency. The expected unit cost for Mr. Smith's two-day rental is $\$8.25 + (\$4.32 \times 2) = \$16.89$, while the expected unit cost for the same rental to Mrs. Jones is $\$16.89 + 0.15 \times (2 \times \$28.00) = \$25.29$. The difference is due to the commission that Rent-a-Lemon will pay to the travel agent if it rents to Mrs. Jones.

The type of activity-based cost model in Table 6.9 is fairly typical for both hotels and rental car companies. The passenger airlines usually estimate a separate incremental cost for each flight leg in their system.

6.7.2 Ancillary Products and Services

It is common in revenue management industries that the fare received from a customer is not the only source of revenue. An airline passenger may also purchase a beer or duty-free goods aboard his flight. Hotels charge for a wide range of goods and services, such as outgoing calls, minibar, and room service. Insurance is a highly profitable sideline for rental car companies. These additional sources of contribution from a customer above and beyond the prices he pays are called *ancillary products or services*.¹⁵ Examples of ancillary products

shows that Mr. Smith always takes out insurance when he rents a car while Ms. Jones never takes out insurance. The net profit from insurance for a two-day rental is \$7.20. Therefore, the net contribution of renting to Mr. Smith would be $2 \times \$28 + \$7.20 - \$16.89 = \$53.51$ and the net contribution of renting to Ms. Jones would be $2 \times \$28 - \$25.29 = \$30.71$.

Note in Example 6.7 that the expected net contribution from renting to Mr. Smith is almost 80% higher than that for Ms. Jones, even though they are paying the same price. If Rent-a-Lemon had only one car available in Topeka, they would be better off renting to Smith than to Jones.

6.8 MEASURING REVENUE MANAGEMENT EFFECTIVENESS

Before deregulation in 1983, airlines could only change their fares after a lengthy regulatory review process. Without direct control over fare, the point of most airline marketing was to fill the planes. The Civil Aeronautics Board always set fares sufficiently high that more passengers meant more money and therefore more profit. Furthermore, there was little or no price differentiation allowed, so there was no scope to perform revenue management. In this environment, the airlines used *load factor* as their primary performance measure. Load factor is defined as the number of seats sold on a flight divided by the number of available seats. A 150-seat flight that departs with 105 passengers on board has experienced a $105/150 = 70\%$ load factor. Load factor was a sensible performance measure in the regulated world. Prices were regulated at a level such that if an airline filled about 75% of its seats on each flight—the so-called *break-even load factor*—they would meet their total costs. Each passenger above the break-even load factor represented additional profit. A marketing program was successful if it increased load factors.

Following deregulation, the issue of performance metrics became more complex. In a multifare world, load factor could no longer pass muster as a leading performance metric. The reason is easy to see—in response to the rise of discount carriers, the major airlines had introduced deep discounts that induced vast amounts of leisure demand on many flights. It would be easy to maximize load factor for these flights—simply allow deep-discount customers to book all the seats. But this would be counter to the entire purpose of tactical revenue management—which is to maximize revenue by reserving seats for late-booking business passengers. Since load factor ignores revenue, it is the wrong metric for a revenue management world. In fact, it is worse than that: Policies designed to maximize load factor would lead to planes filled with deep-discount passengers and massive denial of availability to business passengers—the airlines' highest-paying customers.

To get a better view of the revenue picture, airlines began to look at *yield* in addition to load factor. Yield is simply revenue per passenger mile. Like load factor, it can be measured at any level from flight to market to entire airline. It was sometimes stated that the goal of managing bookings should be to increase yield—hence the term *yield management*, which was the original name for what is now more commonly called *revenue management*. Unfortunately, yield alone is a very imperfect metric since it ignores flight capacity. The strategy for increasing the yield from a flight is the exact opposite from that for increasing

load factor—reject all early-booking demand and accept only late-booking high-yield passengers. This policy would maximize yield but would be disastrous for the airline, since it would result in flights that were empty except for a few high-paying business passengers.

What the airlines really needed was a metric that combined the capacity focus of load factor with the revenue focus of yield and recognized that the goal of revenue management is to maximize the return from resources. For an airline, these resources are seats, and the airline needs to maximize *revenue from available seats*. Since prices for products are at least somewhat proportional to distance, it helps to normalize by distance and to measure *revenue per available seat mile*, or RASM. An airline that flies a 100-seat aircraft on an 1,800-mile flight (about the distance from Chicago to San Francisco) with a total revenue of \$50,000 has just achieved an RASM of $\$50,000 / (100 \times 1,800) = \$0.28$ for that flight. Note that RASM is indifferent to how the revenue was distributed among passengers on the flight. Fifty passengers paying \$1,000 each (a 50% load factor) would result in the same RASM as 100 passengers paying \$500 each (a 100% load factor). This is the right focus from the revenue management point of view—a policy that has resulted in \$51,000 from the flight is more successful than one that resulted in \$50,000, even if the first policy resulted in a lower load factor.¹⁷

Revenue per available seat mile can also be expressed in terms of net yield and load factor:

$$\text{RASM} = \text{Net yield} \times \text{Load factor}$$

Revenue per available seat mile is the leading metric currently used by airlines to measure effectiveness of pricing and revenue management. Similar approaches are also employed at other RM companies: *revenue per available room night* (often abbreviated *REVPAR*) at hotels, *revenue per available rental day* at rental car companies, and *revenue per berth* for cruise lines. The RASM metric enables comparison of revenue performance across markets and within a single market over time. Speaking broadly, flights that are achieving high RASM are generating a higher return to the company's assets than those with low RASM. And RASM also provides a useful performance benchmark among different airlines. Those airlines achieving high RASMs are generating more revenue from their assets in general than those with lower RASMs. However, high RASM does not guarantee high profitability (or any profitability at all), since it is strictly a revenue-based metric and ignores costs entirely.

6.9 REVENUE MANAGEMENT IN ACTION

By most measures, revenue management has been a success. Every major airline, hotel, rental car company, or cruise line has a revenue management department or a revenue management function. These departments rely on computerized systems to recommend actions based on sophisticated mathematical algorithms. The benefits of revenue management at the airlines have been widely acknowledged—so much so that at most major airlines, revenue management is considered a key management function. The practice of revenue management has spread to many other industries, including hotels, rental car companies, cruise lines, railroads, tour operators, broadcasting, and freight transportation. There is active research in many areas. An important area of current research is on how to

better incorporate customer behavior, lifetime customer value, and competitive response into revenue management decisions. This research continues within RM companies, at universities, and at consulting companies and system vendors. In addition, there are numerous efforts to adapt revenue approaches to the needs of new industries, ranging from oil and gas pipelines to health care to made-to-order manufacturing. In each of these industries, constrained perishable capacity plays a major role, and revenue management has the potential to supplement existing pricing approaches, often leading to substantial improvements in profitability.

It would be nice to end the story on this note. On the other hand, as we enter the new century, it seems that the airlines, which are the most sophisticated in revenue management, are among those suffering the most. In early 2003, United Airlines declared bankruptcy, while American Airlines lost more than \$3 billion in 2002 and was asking for federal funding to survive. Meanwhile, the only U.S. airlines making money are the low-cost, low-fare carriers, such as JetBlue and Southwest. It seems to be a complete reversal of the American Airlines/PeopleExpress story—this time it is the carriers with expensive and sophisticated revenue management systems that are going bankrupt while their low-cost, low-fare rivals thrive. What does this say about the future of revenue management?

TABLE 6.11
Comparison of the Big 6 versus low-cost carriers (2002 annual results)

	Average Big 6	Southwest	JetBlue	AirTran	Low-cost average
RASM	10.5	8.4	7.7	9.1	8.4
CASM	11.1	7.5	6.3	8.5	7.4
Net contribution per seat mile	(0.6)	0.9	1.4	0.6	1.0

Some insight into this question is provided by Table 6.11, which shows the average performance of the “Big 6” airlines in the United States (American, Continental, Delta, Northwest, United Airlines, and US Airways) versus three low-cost airlines (Southwest, JetBlue, and AirTran) during 2002.¹⁸ In Table 6.11, CASM stands for *cost per available seat mile*—the *total* operating cost (including fixed operating costs such as crew salaries) divided by the total available seat margin. *Net contribution per seat mile* is simply RASM minus CASM. As can be seen from the table, the Big 6 airlines managed to maintain an average RASM 25% higher than their low-cost competitors. This is an impressive achievement in an environment in which airline seats are rapidly becoming a commodity and the Internet is enabling unprecedented fare visibility. Unquestionably, superior revenue management is a major element enabling the Big 6 to maintain this revenue gap. But while the Big 6 may be winning the revenue game, they are getting clobbered in the cost game. Their average cost per available seat mile is 50% higher than that of the low-cost carriers, overwhelming their revenue advantage. The result is that, in 2002, the low-cost carriers made a penny on average for every seat mile they flew, while the Big 6 lost six-tenths of a cent.

The revenue advantage the Big 6 have been able to maintain is a strong indication of the continuing importance of revenue management. However, it is also unquestionably true

that the game is changing rapidly—not only for airlines but for other RM companies—and revenue management practices are going to need to change as well. Much of this change is being driven by the rise of the Internet as a distribution channel. Retail travel is unquestionably an e-commerce success story: Travel accounted for 36% of retail purchase on the Internet, well ahead of any other category.¹⁹ It is a measure of the success of travel sales on the Internet that some of the leading on-line travel retailers are actually profitable.

From the airlines' point of view, the success of the Internet as a distribution channel has been a mixed blessing. On the one hand, it has enabled the airlines to slash their distribution costs by cutting out the middleman—especially the travel agencies, who used to be the dominant distribution channel for travel products. On the other hand, the Internet has posed a set of new challenges that the major airlines have been struggling to meet.

- The Internet has provided unprecedented fare visibility to consumers. Instead of relying on travel agents or 800 numbers, price-conscious shoppers can surf the net for bargains on their own time, 24 hours a day, seven days a week. This increases the pressure for airlines and others to continually manage their prices and availabilities. It is also eroding some of the traditional “product fences” that have allowed airlines to maintain separate products for leisure and business travelers.
- While the Internet has been a disaster for the traditional travel agency distribution channels, it has led to the rise of new online intermediaries, such as Expedia, Travelocity, and Priceline. These intermediaries and their various competitors are all striving to be the dominant retailer of travel products on the Internet. Naturally, the airlines are leery of the prospect of any intermediary gaining a dominant position, so a consortium of U.S. airlines created Orbitz to be their own online retailer. For the foreseeable future, airlines will need to manage pricing and availability through a variety of online retailers as well as through their own Web sites and the traditional channels.
- The Internet has also created new opportunities for the airlines to create and sell more “inferior” products. For example, Priceline allows customers to bid for travel without knowing the exact departure time or airline they are purchasing. Priceline has explicitly marketed this as an “inferior” product that the airlines can safely sell at a high discount without cannibalizing their mainstream products. Other airlines and hotels are experimenting with selling “distressed inventory” (i.e., empty capacity close to departure) at deep discounts. The need to manage these additional discount products adds further complexity to revenue management.
- As we have seen, revenue management was shaped by the design of the global distribution systems the airlines developed in the 1960s. These systems maintain a relatively small number of fare classes for each product, and the focus of revenue management has been on which of these classes to open and close at different times. Travel agents were trained in the parlance of booking classes and product restrictions that the airlines had established. But the Internet is much more suited to real-time pricing than to availability management. Online shoppers do not care about fare class availabilities; they simply want to find the lowest price for their desired

travel. This means that the entire machinery of protection levels and booking limits behind classical revenue management is becoming less relevant.

While everyone agrees that revenue management needs to change in response to these trends, there is little agreement on what kind of change is needed and how it will occur. But under any scenario, revenue management companies will need to manage a large portfolio of products and market segments through many different channels for some time to come. Until the flexible airplane and hotel is invented, airlines and hotels will need to maximize net contribution from fixed and perishable capacity. The capability to determine who gets to purchase at what fares will be an important driver of market success for some time to come.

It is also clear that the traditional airlines will need to continue to adjust their pricing tactics in response to the challenge of the discount airlines. Already, many of the major airlines have needed to simplify and reduce their fares on routes where they face discount competition. Making a virtue of a necessity, in 2004 Alaska Airlines advertised in their onboard magazine that: "We're proud to announce the retirement of our random & lame-brained fare generator . . ." by "closing the gap between the highest fares and lowest fares in all our markets." Other major airlines were making similar moves, albeit with less fanfare.

Revenue management is in the process of a major paradigm shift. The airlines (and other RM companies) have become experts in managing availabilities. In the new world, they will need to become equally expert in understanding customer price response. Meanwhile, industries outside the travel industry are adopting revenue management approaches to managing availability. The ultimate result may be a convergence of approaches combining revenue management's treatment of availability management with sophisticated customer price-response models.

6.10 SUMMARY

- *Revenue management* is the name given to the ways by which a number of industries maximize expected contribution from their constrained resources. It is applicable to any situation in which:

The seller is selling a fixed stock of perishable capacity.

The seller offers a set of fare classes, each of which has a fixed price (at least in the short run).

The seller can change the availability of fare classes over time.

Customers book capacity prior to usage.

High-fare customers book later than low-fare customers.

Revenue management is used in its purest form by passenger airlines, hotels, rental car companies, cruise lines, passenger trains, apartment rental managers, and various forms of freight transportation. Many of the core concepts are also applicable in other industries.

- Revenue management needs to be implemented at three levels. At the strategic level it requires identifying customer segments and creating "virtual products" and

other differentiated prices. At the booking-control level it requires determining in real time whether or not booking requests should be accepted or rejected. At an intermediate level, tactical revenue management periodically recalculates and updates the booking limits used for booking controls.

- The primary revenue management segmentation at the passenger airlines is between early-booking price-sensitive leisure travelers and later-booking, less price-sensitive business customers. The airlines, hotels, and rental car companies use time-based product differentiation to create products oriented toward each of these segments. In addition, revenue management companies typically have products tailored to other segments, such as groups and large corporations. Finally, they may also use channel pricing and regional pricing to further segment the market. This can lead to pricing structures of great complexity.
- In general, a revenue management company controls a set of constrained resources that can be combined to create products. Each product has a range of discrete fare classes with different associated fares. These fare classes reflect the price differentiation tactics being used by the seller and the set of products and fares the seller has established. Tactical revenue management requires updating booking limits on these fare classes over time.
- Tactical revenue management consists of three interrelated problems:

Capacity allocation—How should fare class booking limits be set for a single-resource product?

Network management—How should bookings for multiresource products be controlled?

Overbooking—How many total bookings should be allowed for a resource?

These three problems are not independent. However, they have different levels of importance within different revenue management industries.

- The main job of a revenue management system is to set initial booking limits and to perform the periodic reoptimization of the booking limits. A revenue management system includes a forecasting module that derives probabilistic forecasts of future demand and an optimization module that uses those forecasts along with other information to determine the optimal booking limits.
- Early revenue management programs aimed to maximize expected revenue. In an environment where incremental costs are low relative to prices, this is a reasonable assumption. However, as prices have fallen and revenue management has been adopted by new industries, it has become increasingly important to estimate incremental costs and include them in the objective function. Further, ancillary products play an important role in many revenue management industries, and their expected contribution needs to be incorporated into revenue management decisions.
- The best measure of overall revenue management performance is contribution per available unit of resource. The metric most typically used at the airlines is

RASM—revenue per available seat mile. Equivalent metrics exist in other revenue management industries. For example, REVPAR—revenue per available room—is the preferred metric for revenue management performance in the hotel industry.

- Revenue management has been a highly effective tactic at the major airlines since the mid-1980s. However, the major airlines are increasingly finding themselves challenged by low-cost carriers. While the major airlines have been able to maintain higher revenues per available seat mile than the low-cost carriers, this revenue advantage has not been able to overcome the cost advantage of the low-cost carriers. It is likely that, at least in the airlines, revenue management will increasingly evolve toward dynamic pricing, with less emphasis on capacity allocation.

The next three chapters address each of the three tactical revenue management problems in more detail. Chapter 7 deals with capacity allocation, Chapter 8 with network management, and Chapter 9 with overbooking.

6.11 EXERCISE

1. A restaurant has 25 tables, each of which can seat up to four people. Typically, it can do three seatings per table during the dinner period from 6:00 to 10:00 PM. The restaurant has a fairly predictable demand pattern, with Fridays and Saturdays the busiest nights and Sundays and Mondays the least busy. The restaurant is always booked to capacity on Fridays and Saturdays and sometimes on Wednesdays and Thursdays but virtually never on Sunday or Monday. It has implemented a revenue management program under which it will sometimes turn down booking requests for one or two people on Fridays and Saturdays if it believes it can fill the restaurant with bookings for larger parties of three and four people. What performance measure should it use to evaluate the effectiveness of its revenue management program?

highest fare class. Then, for the case of normal demand distributions, EMSR-a and EMSR-b can be extended to incorporate buy-up factors as follows:

$$\begin{aligned}\text{EMSR-a with buy-up: } y_j &= \mu_{j-1} + \delta_{j-1} \Phi^{-1} \left[\frac{1}{1-\alpha} \left(1 - \frac{p_j}{p_{j-1}} \right) \right] \\ &\quad - \sum_{i=1}^{j-2} \left[\delta_i \Phi^{-1} \left(1 - \frac{p_j}{p_i} \right) - \mu_i \right] \\ \text{EMSR-b with buy-up: } y_j &= \mu + \delta \Phi^{-1} \left[\frac{1}{1-\alpha} \left(1 - \frac{p_j}{f} \right) \right]\end{aligned}$$

where μ , p , and δ for the modified EMSR-b algorithm are as defined in Equation 7.9.

Modified EMSR-a calculates the protection levels for class j against every higher class under the assumption that buy-up will only occur to the next-highest class, $j-1$, if we close j . It then subtracts the sum of all these protection levels from the remaining capacity to determine the booking limit for class j . EMSR-b assumes that buy-up will occur to the “aggregate class” consisting of all the higher classes and calculates the booking limit using the version of Littlewood’s rule in Equation 7.16.

These two modifications to the EMSR heuristics have been studied by Belobaba and Weatherford (1996) who showed that the modified EMSR heuristics provide additional revenue when buy-up is present. The modified EMSR heuristics are easy to implement, and versions of them have been used by many revenue management companies. However, they have some disadvantages. They require a set of $n-1$ buy-up factors to be estimated and stored for each flight. They also require redefining the demand for each fare class as the demand that would be received if the next-higher class is open. They do not incorporate the possibility of buy-up from class j to classes higher than $j-1$.

Perhaps more fundamentally, the modified EMSR approaches are a “heuristic grafted on to another heuristic.” They are not derived from a fundamental model of consumer choice. Rather, they assume that passengers have M-Class and Y-Class stamped on their heads and that only some of the M-Class passengers will buy up to Y-Class if M-Class is closed. Development of more fundamental models including consumer choice is an active research area.

7.5 CAPACITY ALLOCATION IN ACTION

We have described several techniques for calculating booking limits. These techniques need to be used within the context of a booking control structure, as described in Section 6.5. In practice, booking limits are periodically calculated offline for a flight and transmitted to the reservation system where they are used to control bookings. The combination of periodic recalculation of booking limits and the nested booking limits within the reservation system defines a *dynamic booking control process*.

Dynamic Booking Control Process

1. Calculate the initial booking limits for a flight.
2. When a booking request for class j is received, compare the number of seats requested with the current booking limit for the class.

3. If the number of seats requested exceeds the booking limit for class j , reject the request and go to step 6.
4. If the number of seats requested is within the booking limit for class j , accept the request. Depending on the capabilities of the reservation system, either (a) decrement the booking limits for classes with positive availability by the number of seats accepted or (b) decrement the booking limits for *all* classes by the number of seats accepted.
5. When a booking cancels, either (a) increment the booking limits for classes with positive availability by the number of seats cancelled (irreversible) or (b) increment to booking limits for *all* classes by the number of seats cancelled (reversible).
6. Periodically recalculate all booking limits based on total bookings and demand expected in each class until departure. (This recalculation is often called a *reoptimization*.)

This dynamic booking control process makes no assumption about the nature of the re-optimization approach being used or the frequency with which reoptimizations occur. And, in fact, there is considerable variation among companies in the optimization approaches used and the frequency and situations under which they are applied. There is a general belief that frequent reoptimization is important—particularly when demand is changing rapidly or departure is approaching and bookings and cancellations are being received at a high rate. Often, the tradeoff is between using a more complex algorithm and reoptimizing less often and using a simpler algorithm and reoptimizing more often. These issues are the topics of ongoing research and are by no means settled.

7.6 MEASURING CAPACITY ALLOCATION EFFECTIVENESS

We saw in Chapter 6 that revenue per available seat mile (RASM) is the metric most widely used to measure the overall effectiveness of revenue management at an airline. RASM is a good high-level measure, but it does not tell the entire story. It is often not immediately clear why RASM is high or low in a particular market. Are we achieving a high RASM in a particular market relative to the competition because we have superior service? superior marketing? better revenue management? all of these? RASM by itself does not address the important need to isolate and measure the effects of revenue management separately from all other influences in a market. The need to understand the effects of revenue management is important because revenue management software systems are expensive, usually requiring millions of dollars of investment in software and hardware. Revenue management programs are also expensive to operate—a major airline may have a department of 50 or more people dedicated to running the revenue management system and reviewing and updating its forecasts and recommendations. How can management be convinced that this expensive and somewhat esoteric department is yielding results that justify this investment? A metric specifically targeted at measuring revenue management effectiveness is required.

The most popular approach to measuring revenue management effectiveness is the *rev-*

11.1 BACKGROUND AND BUSINESS SETTING

All of the pricing tactics we have studied so far assume that buyers and sellers interact in the same basic fashion. The seller offers a stock of goods to different customer segments through different channels. Potential customers arrive, observe a price, and decide whether or not to purchase. The seller needs to decide what price to offer for each product to each customer group through each channel. The seller monitors sales of his goods and periodically updates the prices and/or availabilities he is offering. This is the *list-pricing* model. All of the pricing settings we have covered so far, including peak load pricing, revenue management, and markdown management, are varieties of list pricing.

List pricing may seem quite general. However, there is another important class of pricing—*customized pricing*—that is of equal or greater importance in many industries. In customized pricing, potential customers approach the seller, one by one, and the seller quotes each one a price. Often (but not always) each potential buyer wants something different—either a different bundle of products or services or a different quantity or a different variation on a basic product. In customized pricing, the seller can quote a different price for each request. The price can be based on knowledge of the individual buyer, the product(s) requested by the buyer, and other factors, such as general market conditions and competitive offerings. We now describe some examples of customized pricing.

11.1.1 Telecommunication Services

A company sells telecommunications services in the North American market. The company has a list of more than 40 different services it provides, including local, long-distance, and international voice transmission as well as data communication and corporate networking services. Different customers request different combinations of these services. Services are sold on a contract basis, with customers typically agreeing to a contract under which the company provides all of their telecommunication services for the next year. The company has three major competitors who can provide the same spectrum of services it offers, as well

UNIVERSITY OF WASHINGTON LIBRARIES
All overdue materials are subject to fines.
Due date may be changed if needed by others.
DATE DUE

RECEIVED SEP 02 2005
RECEIVED SEP 21 2005

MAR 17 2006

MAR 17 2006

Field Business

MAY 17 2006

MAY 30 2007

Field Business

[Signature]

[Signature]

UWB/CCC AUG 29 2007 12 WEEK

FOSTER BUSINESS LIBRARY
BOX 353224
University of Washington
Seattle, WA 98195-3224



3 9352 06503735 0

HF5416.5 .P457 2005

Foster General Stacks

PHILLIPS

PRICING AND REVENUE
OPTIMIZATION

HF

5416.5

P457

2005

